

**Cengage Advantage**  
Books

NINTH EDITION

# Understanding Arguments

AN INTRODUCTION TO INFORMAL LOGIC

**Walter Sinnott-Armstrong  
Robert Fogelin**

# **U N D E R S T A N D I N G**

# **A R G U M E N T S**

---



# **UNDERSTANDING ARGUMENTS**

---

*An Introduction to Informal Logic*

**NINTH EDITION**

**WALTER SINNOTT-ARMSTRONG**

*Duke University*

**ROBERT J. FOGELIN**

*Dartmouth College*



---

Australia • Brazil • Mexico • Singapore • United Kingdom • United States

This is an electronic version of the print textbook. Due to electronic rights restrictions, some third party content may be suppressed. Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. The publisher reserves the right to remove content from this title at any time if subsequent rights restrictions require it. For valuable information on pricing, previous editions, changes to current editions, and alternate formats, please visit [www.cengage.com/highered](http://www.cengage.com/highered) to search by ISBN#, author, title, or keyword for materials in your areas of interest.

**Understanding Arguments: An Introduction to Informal Logic, Ninth Edition****Walter Sinnott-Armstrong and Robert Fogelin**

Product Manager: Debra Matteson

Content Developer: Ian Lague

Content Coordinator: Joshua Duncan

Media Developer: Christian Biagetti

Associate Marketing Manager: Shanna Shelton

Rights Acquisitions Specialist: Shalice Shah-Caldwell

Manufacturing Planner: Sandee Milewski

Art and Design Direction, Production Management, and Composition: PreMediaGlobal

Cover Image: © DEA PICTURE LIBRARY/De Agostini Picture Library / Getty Images

© 2015, 2010, 2005 Cengage Learning

WCN: 02-200-203

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher.

For product information and technology assistance, contact us at **Cengage Learning Customer & Sales Support, 1-800-354-9706**

For permission to use material from this text or product, submit all requests online at [www.cengage.com/permissions](http://www.cengage.com/permissions).

Further permissions questions can be emailed to  
[permissionrequest@cengage.com](mailto:permissionrequest@cengage.com).

Library of Congress Control Number: 2013952536

Student Edition:

ISBN-13: 978-1-285-19736-4

ISBN-10: 1-285-19736-4

**Cengage Learning**

200 First Stamford Place, 4th Floor  
Stamford, CT 06902  
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil and Japan. Locate your local office at [international.cengage.com/region](http://international.cengage.com/region).

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

For your course and learning solutions, visit [www.cengage.com](http://www.cengage.com).

Purchase any of our products at your local college store or at our preferred online store [www.cengagebrain.com](http://www.cengagebrain.com).

**Instructors:** Please visit [login.cengage.com](http://login.cengage.com) and log in to access instructor-specific resources.

Printed in the United States of America

1 2 3 4 5 6 7 17 16 15 14 13

*To Eric, John, Lars, Miranda, and Nicholas  
and the colleges of their choice*

---



---

**CONTENTS**

---

**PREFACE XV****PART I HOW TO ANALYZE ARGUMENTS 1**

---

**CHAPTER 1****USES OF ARGUMENTS 3**

<i>What Arguments Are</i>	3
<i>Justifications</i>	4
<i>Explanations</i>	7
<i>Combinations: An Example</i>	10

**CHAPTER 2****THE WEB OF LANGUAGE 17**

<i>Language and Convention</i>	17
<i>Linguistic Acts</i>	19
<i>Speech Acts</i>	22
Performatives	23
Kinds of Speech Acts	26
<i>Conversational Acts</i>	28
Conversational Rules	31
Conversational Implication	33
Rhetorical Devices	36
<i>Summary</i>	38

**CHAPTER 3****THE LANGUAGE OF ARGUMENT 41**

<i>Argument Markers</i>	41
If . . . , then . . .	43
<i>Arguments in Standard Form</i>	45
<i>A Problem and Some Solutions</i>	47
Assuring	48

Guarding	49
Discounting	51
<i>Evaluative Language</i>	54

**CHAPTER 4**  
**THE ART OF CLOSE ANALYSIS 59**

<i>An Extended Example</i>	59
Clerk Hire Allowance, House of Representatives	59

**CHAPTER 5**  
**DEEP ANALYSIS 79**

<i>Getting Down to Basics</i>	79
<i>Clarifying Crucial Terms</i>	83
<i>Dissecting the Argument</i>	83
<i>Arranging Subarguments</i>	85
<i>Some Standards for Evaluating Arguments</i>	90
Validity	91
Truth	93
Soundness	94
<i>Suppressed Premises</i>	96
Contingent Facts	97
Linguistic Principles	99
Evaluative Suppressed Premises	100
Uses and Abuses of Suppressed Premises	100
<i>The Method of Reconstruction</i>	102
<i>An Example of Reconstruction: Capital Punishment</i>	105

---

**PART II HOW TO EVALUATE ARGUMENTS: DEDUCTIVE STANDARDS 111**

---

**CHAPTER 6**  
**PROPOSITIONAL LOGIC 113**

<i>The Formal Analysis of Arguments</i>	113
<i>Basic Propositional Connectives</i>	114
Conjunction	114
Disjunction	122
Negation	122
Process of Elimination	125
How Truth-Functional Connectives Work	126

Testing for Validity	128
Some Further Connectives	132
<i>Conditionals</i>	134
Truth Tables for Conditionals	136
Logical Language and Everyday Language	142
Other Conditionals in Ordinary Language	145

**CHAPTER 7****CATEGORICAL LOGIC 151**

<i>Beyond Propositional Logic</i>	151
<i>Categorical Propositions</i>	152
The Four Basic Categorical Forms	154
Translation into the Basic Categorical Forms	156
Contradictries	159
Existential Commitment	161
<i>Validity for Categorical Arguments</i>	162
Categorical Immediate Inferences	164
The Theory of the Syllogism	166

---

**PART III HOW TO EVALUATE ARGUMENTS: INDUCTIVE STANDARDS 177****CHAPTER 8****ARGUMENTS TO AND FROM GENERALIZATIONS 179**

<i>Induction versus Deduction</i>	179
<i>Statistical Generalizations</i>	183
Should We Accept the Premises?	184
Is the Sample Large Enough?	185
Is the Sample Biased?	186
Is the Sampling Procedure Biased?	187
<i>Statistical Applications</i>	189

**CHAPTER 9****INFERENCE TO THE BEST EXPLANATION AND FROM ANALOGY 195**

<i>Inferences to the Best Explanation</i>	195
Which Explanation Is Best?	198
Context Is Crucial	200
<i>Arguments from Analogy</i>	204
Are Analogies Explanations?	207

**CHAPTER 10****CAUSAL REASONING 215**

<i>Reasoning About Causes</i>	215
<i>Sufficient Conditions and Necessary Conditions</i>	217
The Sufficient Condition Test	220
The Necessary Condition Test	221
The Joint Test	222
Rigorous Testing	225
Reaching Positive Conclusions	226
<i>Applying These Methods to Find Causes</i>	228
Normality	228
Background Assumptions	229
A Detailed Example	230
<i>Concomitant Variation</i>	234

**CHAPTER 11****CHANCES 239**

<i>Some Fallacies of Probability</i>	239
The Gambler's Fallacy	239
Heuristics	241
<i>The Language of Probability</i>	243
<i>A Priori Probability</i>	244
<i>Some Rules of Probability</i>	246
Probabilities of Negations	246
Probabilities of Conjunctions	247
Probabilities of Disjunctions	248
Probabilities in a Series	249
Permutations and Combinations	250
<i>Bayes's Theorem</i>	253

**CHAPTER 12****CHOICES 263**

<i>Expected Monetary Value</i>	263
<i>Expected Overall Value</i>	266
<i>Decisions Under Ignorance</i>	268

**PART IV FALLACIES 273****CHAPTER 13****FALLACIES OF VAGUENESS 275**

- Uses of Unclarity* 275  
*Vagueness* 276  
*Heaps* 278  
*Slippery Slopes* 280  
    Conceptual Slippery-Slope Arguments 280  
    Fairness Slippery-Slope Arguments 282  
    Causal Slippery-Slope Arguments 285

**CHAPTER 14****FALLACIES OF AMBIGUITY 291**

- Ambiguity* 291  
*Equivocation* 295  
*Definitions* 299

**CHAPTER 15****FALLACIES OF RELEVANCE 307**

- Relevance* 307  
*Ad Hominem Arguments* 308  
    Inconsistency 311  
    Genetic Fallacies 311  
*Appeals to Authority* 314  
*More Fallacies of Relevance* 318  
    Appeals to Popular Opinion 318  
    Appeals to Emotion 319

**CHAPTER 16****FALLACIES OF VACUITY 323**

- Circularity* 323  
*Begging the Question* 324  
*Self-Sealers* 328

**CHAPTER 17  
REFUTATION 333**

- What Is Refutation?* 333  
*Counterexamples* 334  
*Reductio Ad Absurdum* 337  
*Straw Men and False Dichotomies* 341  
*Refutation by Parallel Reasoning* 343

**PART V AREAS OF ARGUMENTATION 351****CHAPTER 18  
LEGAL REASONING 353**

- Components of Legal Reasoning* 354  
Questions of Fact 354  
Questions of Law 355  
*The Law of Discrimination* 361  
The Equal Protection Clause 361  
Applying the Equal Protection Clause 362  
The Strict Scrutiny Test 363  
The *Bakke* Case 364  
Regents of the University of California v. Bakke 366  
Legal Developments Since *Bakke* 368  
*Grutter v. Bollinger* 369  
*Gratz v. Bollinger* 375  
*Burden of Proof* 381

**CHAPTER 19  
MORAL REASONING 383**

- Moral Disagreements* 383  
*The Problem of Abortion* 384  
The “Pro-Life” Argument 385  
“Pro-Choice” Responses 387  
*Analogical Reasoning in Ethics* 392  
*Weighing Factors* 394  
“Abortion,” by Mary Anne Warren 397  
“An Argument that Abortion Is Wrong,”  
by Don Marquis 409

**CHAPTER 20****SCIENTIFIC REASONING 423***Standard Science* 423*Scientific Revolutions* 425

“Molecular Machines: Experimental Support for the Design

Inference,” by Michael J. Behe 427

“Living with Darwin,” by Philip Kitcher 440

**CHAPTER 21****RELIGIOUS REASONING 449**

“Five Reasons to Believe in God,” by William Lane Craig 450

“Seven Deadly Objections to Belief in the Christian God,”  
by Edwin Curley 456**CHAPTER 22****PHILOSOPHICAL REASONING 465**“Sanity and the Metaphysics of Responsibility,”  
by Susan Wolf 469“A Defence of Free Will Skepticism,”  
by Derk Pereboom 483**INDEX 495**



---

**P R E F A C E**

---

Traditionally, logic has been considered the most general science dealing with arguments. The task of logic is to discover the fundamental principles for distinguishing good arguments from bad ones.

For certain purposes, arguments are best studied as abstract patterns of reasoning. Logic can then focus on these general forms rather than on particular arguments, such as your attempt to prove to the bank that they, not you, made a mistake. The study of those general principles that make certain patterns of argument valid and other patterns of argument invalid is called *formal logic*. Two chapters of this work are dedicated to formal logic.

A different but complementary way of viewing an argument is to treat it as a particular use of language: Presenting arguments is one of the important things we do with words. This approach stresses that arguing is a linguistic activity. Instead of studying arguments as abstract patterns, it examines them as they occur in concrete settings. It raises questions of the following kind:

What is the place of argument within language as a whole?

What words or phrases are characteristic of arguments?

How do these words function?

What task or tasks are arguments supposed to perform?

When an approach to argument has this emphasis, the study is called *informal logic*. Though it contains a substantial treatment of formal logic, *Understanding Arguments*, as its subtitle indicates, is primarily a textbook in informal logic.

The ninth edition of *Understanding Arguments* differs from the eighth edition in a number of ways. The most important change is simplification. Many chapters have been shortened and streamlined. Our goal was to remove tangents and complexities that confuse students so that the main points could be understood more easily. In addition, the different kinds of inductive arguments have been reordered to provide a better flow between topics. Several sections have been split up and reorganized for clarity. Some of the more difficult and confusing topics have been dropped. This edition also contains new readings on moral and philosophical reasoning in Chapters 19 and 22. These new readings make the text more accessible and relevant to popular debates. Finally, we updated many examples, exercises, and discussion questions throughout the text.

Another major innovation that begins with this Ninth Edition is that readers of *Understanding Arguments* will also have free access to detailed lectures, exercises, and quizzes on all of the material in Chapters 1–17. There are over 100 mini-lectures keyed to almost 1000 corresponding questions. These supplementary materials are available for free on the Coursera website (<https://www.coursera.org/>) in a MOOC titled Think Again: How to Reason and Argue. We hope that these lectures and exercises help readers get the most out of this book.

This new edition has been influenced by our teaching of this material with various colleagues. In this regard, we would especially like to thank Ram Neta, who co-taught with Walter Sinnott-Armstrong on the Coursera website, as well as to many people at Coursera and in the Office of Instructional Technology at Duke University, who taught us to make these topics more accessible and lively. We received invaluable help from two student assistants—Joe Metz and Jason Bowers—in addition to the many others who helped us on previous editions. We are also indebted to the following reviewers: Dan Berger, Simpson University; William Brunson, University of Nevada—Las Vegas; Aaron Cobb, Auburn University—Montgomery; Nathaniel Goldberg, Washington and Lee University; Deke Gould, Augustana College; Robert Bruce Kelsey, Thomas College; Jung Kwon, Long Beach City College; Judith Little, SUNY—Potsdam; Diane Michelfelder, Macalester College; Rachel Mohr, University of Nebraska—Lincoln; Dennis Montgomery, Norfolk State University; Patrice Nango, Mesa Community College; Kurt Nutting, San Francisco State University; Michael Patton, University of Montevallo; Marc Pugliese, Saint Leo University; Eric Rovie, Georgia Perimeter College—Newton; and Catherine Womack, Bridgewater State University. At Cengage Learning and PreMediaGlobal, we received expert advice and assistance from Joann Kozyrev—Senior Sponsoring Editor, Debra Matteson—Product Manager, Prashanth Kamavarapu—Project Manager, Ian Lague—Development Editor, Kristina Mose-Libon—Art Director, and Joshua Duncan—Assistant Editor. Without all of these people, this book would contain many more mistakes than it undoubtedly still does.

Walter Sinnott-Armstrong  
Robert J. Fogelin

# HOW TO ANALYZE ARGUMENTS

*Arguments are all around us. They bombard us constantly in advertisements; in courtrooms; in political, moral, and religious debates; in academic courses on mathematics, science, history, literature, and philosophy; and in our personal lives when we make decisions about our careers, finances, and families. These crucial aspects of our lives cannot be understood fully without understanding arguments. The goal of this book, then, is to help us understand arguments and, thereby, to understand our lives.*

*We will view arguments as tools. To understand a tool, we need to know the purposes for which it is used, the material out of which it is made, and the forms that it takes. For example, hammers are normally used to drive nails or to pound malleable substances. Hammers are usually made out of a metal head and a handle of wood, plastic, or metal. A typical hammer's handle is long and thin, and its head is perpendicular to its handle. Similarly, in order to understand arguments, we need to investigate their purposes, materials, and forms.*

*Chapter 1 discusses the main purposes or uses of arguments. The material from which arguments are made is language, so Chapters 2 and 3 explore language in general and then the language of argument in particular. Chapters 4 and 5 use the lessons learned by then to analyze concrete examples of arguments in detail. The following chapters turn to the forms of arguments, including deductive forms in Part II (Chapters 6 and 7) and inductive forms in Part III (Chapters 8–12). Each form of argument comes with its own standards of adequacy. Part IV (Chapters 13–17) will then consider the main ways in which arguments can go astray, including fallacies of clarity, relevance, and vacuity. By the end of this journey, we should understand arguments much better.*



## USES OF ARGUMENTS

**What are arguments?** In our view, arguments are tools, so the first step toward understanding arguments is to ask what they are used for—what people are trying to accomplish when they give arguments. This brief chapter will propose a definition of arguments and then explore two main purposes of arguments: justification and explanation. Both justifications and explanations try to provide reasons, but reasons of different kinds. Justifications are supposed to give reasons to believe their conclusions, whereas explanations are supposed to give reasons why their conclusions are true. Each of these purposes is more complicated and fascinating than is usually assumed.

### WHAT ARGUMENTS ARE

---

The word “argument” may suggest quarrels or squabbles. That is what a child means when she reports that her parents are having an argument. Arguments of that sort often include abuse, name-calling, and yelling. That is not what this book is about. The goal here is not to teach you to yell louder, to be more abusive, or to beat your opponents into submission.

Our topic is the kind of argument defined by Monty Python in their justly famous “Argument Clinic.” In this skit, a client enters a clinic and pays for an argument. In the first room, however, all he gets is abuse, which is not argument. When he finally finds the right room to get an argument, the person who is supposed to give him an argument simply denies whatever the client says, so the client complains that mere denial is different from argument, because “an argument is a connected series of statements to establish a definite proposition.” This definition is almost correct. As we will see, the purpose of an argument need not always be to “establish” its conclusion, both because some conclusions were established in advance and because many reasons are inconclusive. Nonetheless, Monty Python’s definition needs to be modified only a little in order to arrive at an adequate definition:

An argument is a connected series of sentences, statements, or propositions (called “premises”) that are intended to give a reason of some kind for a sentence, statement, or proposition (called the “conclusion”).

This definition does not pretend to be precise, but it does tell us what arguments are made of (sentences, statements, or propositions) and what their purpose is (to give reasons).

Another virtue of this definition is that it is flexible enough to cover the wide variety of arguments that people actually give. Different arguments are intended to give reasons of very different sorts. These reasons might be justificatory reasons to believe or to disbelieve some claim. They might, instead, be explanatory reasons why something happened. They might even be practical reasons to do some act. Because reasons come in so many kinds, arguments are useful in a great variety of situations in daily life. Trying to determine why your computer crashed, why your friend acted the way she did, and whether it will rain tomorrow as well as trying to decide which political candidate to vote for, which play to use at a crucial point in a football game, where to go to college, and whether to support or oppose capital punishment—all involve weighing and evaluating reasons.

It is inaccurate, therefore, to think of arguments as serving only one single, simple purpose. People often assume that you always use every argument to make other people believe what you believe and what they did not believe before hearing or reading the argument. Actually, however, some arguments are used for that purpose, but others are not. To fully understand arguments in all their glory, then, we need to distinguish different uses of argument. In particular, we will focus on two exemplary purposes: justification and explanation.

## JUSTIFICATIONS

---

One of the most prominent uses of arguments is to justify a disputed claim. For example, if I claim that September 11, 2001, was a Tuesday, and you deny this or simply express some doubt, then we might look for a calendar. But suppose we don't have a calendar for 2001. Luckily, we do find a calendar for 2002. Now I can justify my claim to you by presenting this argument: The calendar shows that September 11 was on Wednesday in 2002; 2002 was not a leap year, since 2002 is not divisible by 4; nonleap years have 365 days, which is 1 more day than 52 weeks; so September 11 must have been on Tuesday in 2001. You should now be convinced.

What have I done? My utterance of this argument has the *effect* of changing your mind by getting you to believe a conclusion that you did not believe before. Of course, I might also be able to change your mind by hypnotizing you. But normally I do not want to use hypnosis. I also do not want to change your mind by manufacturing a fake calendar for 2002 with the wrong dates or by fooling you with a bad argument. Such tricks would not satisfy my goals fully. This shows that changing your mind is not all that I am trying to accomplish. I want more than simply to *persuade* you or *convince*

you. What else do I want? My additional aim is to show you that you *should* change your mind, and why. I want my argument to be *good* and to give you a *good reason* to change your mind. I want my argument not only to persuade you but also to make you *justified* in believing my conclusion.

The above example is typical of one kind of justification, but there are other patterns. Suppose that I share your doubts about which day of the week it was on September 11, 2001. Then I might use the same argument to justify my belief as well as yours. Indeed, you don't even need to be present. If I am all alone, and I just want to figure out which day of the week it was on September 11, 2001, then I might think in terms of this same argument. Here the goal is not to convince anybody else, but the argument is still used to find a good reason to believe the conclusion.

In cases like these, we can say that the argument is used for *impersonal normative justification*. The justification is normative because the goal is to find a reason that is a good reason. It is impersonal because what is sought is a reason that is or should be accepted as a good reason by everyone capable of grasping this argument, regardless of who they are. The purpose is to show that there *is* a reason to believe the conclusion, regardless of who *has* a reason to believe it. Other arguments, in contrast, are aimed at specific people, and the goal is to show that those particular people are committed to the conclusion or have a reason to believe the conclusion. Such individualized uses of arguments seek what can be called *personal justification*.

There should be nothing surprising about different people having different reasons. I might climb a mountain to appreciate the view at the top, whereas you climb it to get exercise, and your friend climbs it to be able to talk to you while you climb it. Different people can have different reasons for the same action. Similarly, different people can have different reasons to believe the same conclusion. Suppose that someone is murdered in the ballroom with a revolver. I might have good reason to believe that Miss Peacock did not commit the murder, because I saw her in the library at the time the murder was committed. You might not trust me when I tell you that I saw her, but you still might have good reason to believe that she is innocent, because you believe that Colonel Mustard did it alone. Even if I doubt that Colonel Mustard did it, we still each have our own reasons to agree that Miss Peacock is innocent.

When different people with different beliefs are involved, we need to ask who is supposed to accept the reason that is given in an argument. A speaker might give an argument to show a listener that the speaker has a reason to believe something, even though the speaker knows that the audience does not and need not accept that reason. Suppose that you are an atheist, but I am an evangelical Christian, and you ask me why I believe that Jesus rose from the dead. I might respond that the Bible says that Jesus rose from the dead, and what the Bible says must be true, so Jesus rose from the dead. This argument tells you what *my* reasons are for believing what I believe, even if *you* do not accept those reasons. My argument can be used to show you that

I have reasons and what my reasons are, regardless of whether you believe that my reasons are good ones and also regardless of whether my reasons really are good ones.

The reverse can also happen. A speaker might give an argument to show a listener that the listener has a reason to believe something, even though the speaker does not accept that reason. Suppose that you often throw loud parties late into the night close to my bedroom. I want to convince you to stop or at least quiet down. Fortunately, you think that every citizen ought to obey the law. I disagree, for I am an anarchist bent on undermining all governments and laws. Still, I want to get a good night's sleep before the protest tomorrow, so I might argue that it is illegal to make that much noise so late, and you ought to obey the law, so you ought to stop throwing such loud parties. This argument can show you that *you* are committed to its conclusion, even if *I* believe that its premises are false.

Of course, whether I succeed in showing my audience that they have a reason to believe my conclusion depends on who my audience is. My argument won't work against loud neighbors who don't care about the law. Consequently, we need to know who the audience is and what they believe in order to be able to show them what reason they have to believe a conclusion.

In all of these cases, arguments are used to show that someone has a reason to believe the conclusion of the argument. That is why all of these uses can be seen as providing different kinds of justification. The differences become crucial when we try to evaluate such arguments. If my goal is to show you that you have a reason to believe something, then I can be criticized for using a premise that you reject. Your beliefs are no basis for criticism, however, if all I want is to show my own reasons for believing the conclusion. Thus, to evaluate an argument properly, we often need to determine not only whether the argument is being used to justify a belief but also which kind of justification is sought and who the audience is.

### EXERCISE 1

Write the best brief argument you can to justify each of the following claims to someone who does not believe them.

1. Nine is not a prime number.
2. Seven is a prime number.
3. A molecule of water has three atoms in it.
4. Water is not made up of carbon.
5. The U.S. president lives in Washington, D.C.
6. The Earth is not flat.
7. Humans have walked on the moon.
8. Most bicycles have two wheels.

**DISCUSSION QUESTION**

When, if ever, is it legitimate to try to convince someone else to believe something on the basis of a premise that you yourself reject? Consider a variety of cases.

**EXPLANATIONS**

A different but equally important use of arguments is to provide explanations. Explanations answer questions about how or why something happened. We explain how a mongoose got out of his cage by pointing to a hole he dug under the fence. We explain why Smith was acquitted by saying that he got off on a technicality. The purpose of explanations is not to prove that something happened, but to make sense of things.

An example will bring out the difference between justification and explanation. One person claims that a school's flagpole is thirty-five feet tall, and someone else asks her to justify this claim. In response, she might produce a receipt from the Allegiance Flagpole Company acknowledging payment for a flagpole thirty-five feet in height. Alternatively, she may put a stick straight up into the ground, measure the stick's length and its shadow's length, then measure the length of the flagpole's shadow, and calculate the length of the flagpole. Neither of these justifications, however, will answer a different question: *Why* is the flagpole thirty-five feet tall? This new question could be answered in all sorts of ways, depending on context: The school could not afford a taller one. It struck the committee as about the right height for the location. That was the only size flagpole in stock. There is a state law limiting flagpoles to thirty-five feet. And so on. These answers help us understand why the flagpole is thirty-five feet tall. They explain its height.

Sometimes simply filling in the details of a story provides an explanation. For example, we can explain how a two-year-old girl foiled a bank robbery by saying that the robber tripped over her while fleeing from the bank. Here we have made sense out of an unusual event by putting it in the context of a plausible *narrative*. It is unusual for a two-year-old girl to foil a bank robbery, but there is nothing unusual about a person tripping over a child when running recklessly at full speed in a crowded area.

Although the narrative is probably the most common form of explanation in everyday life, we also often use arguments to give explanations. We can explain a certain event by deriving it from established principles and accepted facts. This argument then has the following form:

- (1) General principles or laws
- (2) A statement of initial conditions
- ∴ (3) A statement of the phenomenon to be explained

The symbol “∴” is pronounced “therefore” and indicates that the premises above the line are supposed to give a reason for the conclusion below the

line. By “initial conditions” we mean those facts in the context that, together with appropriate general principles and laws, allow us to derive the result that the event to be explained occurs.

This sounds quite abstract, but an example should clarify the basic idea. Suppose we put an ice cube into a glass and then fill the glass with water to the brim. The ice will stick out above the surface of the water. What will happen when the ice cube melts? Will the water overflow? Will it remain at the same level? Will it go down? Here we are asking for a *prediction*, and it will, of course, make sense to ask a person to *justify* whatever prediction he or she makes. Stumped by this question, we let the ice cube melt to see what happens. We observe that the water level remains unchanged. After a few experiments, we convince ourselves that this result always occurs. We now have a new question: *Why* does this occur? Now we want an explanation of this phenomenon. The explanation turns upon the law of buoyancy, which says that an object in water is buoyed up by a force equal to the weight of the water it displaces. This law implies that, if we put an object in water, it will continue to sink until it displaces a volume of water whose weight is equal to its own weight (or else the object hits the bottom of the container). With this in mind, go back to the original problem. An ice cube is itself simply water in a solid state. Thus, when it melts, it will exactly fill in the volume of water it displaced, so the water level will remain unchanged.

We can now see how this explanation conforms to the argumentative pattern mentioned above:

- (1) General principles or laws (Primarily the law of buoyancy)
- (2) Initial conditions (An ice cube in a glass of water filled to the brim)
- . . . (3) Phenomenon explained (The level of the water remaining unchanged after the ice cube melts)

This explanation is fairly good. People with only a slight understanding of science can follow it and see why the water level remains unchanged. We should also notice that it is not a *complete* explanation, because certain things are simply taken for granted—for example, that things do not change weight when they pass from a solid to a liquid state. To put the explanation into perfect argumentative form, this assumption and many others would have to be stated explicitly. This is never done in everyday life and is only rarely done in the most exact sciences.

Is this explanation any good? Explanations are satisfactory if they remove bewilderment or surprise by telling us *how* or *why* something happened in a way that is relevant to the concerns of a particular context. Our example does seem to accomplish that much. However, it might seem that even the best explanations are not very useful because they take so much for granted. In explaining why the water level remains the same when the ice cube melts, we cited the law of buoyancy. Now, why should that law be true? What explains *it*? To explain the law of buoyancy, we would have to derive it from other laws that are more general and, perhaps, more intelligible. In fact, this

has been done. Archimedes simultaneously proved and explained the law of buoyancy by deriving it from the laws of the lever. How about the laws of the lever? Can they be proved and explained by deriving them from still higher and more comprehensive laws? Perhaps. Yet reasons give out, and sooner or later explanation (like justification) comes to an end. It is the task of science and all rational inquiry to move that boundary further and further back. But even when there is more to explain, that does not show that a partial explanation is totally useless. As we have seen, explanations can be useful even when they are incomplete, and even though they are not used to justify any disputed claim. Explanation is, thus, a separate use of arguments.

### EXERCISE II

Houses in Indonesia sometimes have their electrical outlets in the middle of the wall rather than at floor level. Why? A beginning of an explanation is that flooding is a danger in the Netherlands. Citing this fact does not help much, however, unless one remembers that Indonesia was formerly a Dutch colony. We can understand why the Dutch might put their electrical outlets above floor level in the Netherlands. It is safer in a country where flooding is a danger. Is flooding, then, a similar danger in Indonesia? Apparently not; so why did the Dutch continue this practice in Indonesia? The answer is that colonial settlers tend to preserve their home customs, practices, and styles. The Dutch continued to build Dutch-style houses with the electrical outlets where (for them) they are normally placed—that is, in the middle of the wall rather than at floor level. Restate this explanation in the form of an argument (that is, specify its premises and conclusion).

### EXERCISE III

Write a brief argument to explain each of the following. Indicate what facts and what general principles are employed in your explanations. (Do not forget those principles that may seem too obvious to mention.)

1. Why a lighter-than-air balloon rises.
2. Why there is an international date line.
3. Why average temperatures tend to be higher closer to the equator.
4. Why there are usually more college freshmen who plan to go to medical school than there are seniors who still plan to go to medical school.
5. Why almost no textbooks are more than eighteen inches high.
6. Why most cars have four tires (instead of more or fewer).
7. Why paintings by Van Gogh cost so much.
8. Why wages go up when unemployment goes down.

**DISCUSSION QUESTION**

It is sometimes said that science tells us *how* things happen but does not tell us *why* they happen. In what ways is this contention right, and in what ways is it wrong?

**COMBINATIONS: AN EXAMPLE**

Although justification and explanation are distinct uses of arguments, we often want to know both *what* happened and also *why* it happened. Then we need to combine justifications and explanations. We can see how this works by considering a fictional crime.

Imagine that Madison was arrested for murdering her husband, Victor. Now she is on trial, and you are on the jury. Presumably, the police and the prosecuting attorneys would not have arrested and prosecuted her if they did not believe that Madison committed the murder, but are their beliefs justified? Should she be convicted and sent to prison? That's up to you and the other jurors to decide.

You do not want to convict her arbitrarily, of course, so you need arguments to justify you in believing that Madison is guilty. The goal of prosecuting attorneys is to provide such justification. Their means of reaching this goal is to present evidence and arguments during the trial. Although their ultimate conclusion is that you should find Madison guilty of murder, the prosecutors need to justify lots of little claims along the way.

It might seem too obvious to mention, but the prosecution first needs an argument to show that the victim died. After all, if nobody died, nobody was killed. This first argument can be pretty simple: This person was walking and talking before he was shot in the head; now his heart has stopped beating for a long time; so he must be dead. There can be complications, since some gunshot victims can be revived, but let's assume that an argument like this justifies the claim that the victim is dead.

We also want to know who the victim was. The body was identified by several of Victor's friends, we assume, so all the prosecution needs to argue is that identifications like this are usually correct, so it was Victor who died. This second argument also provides a justification, but it differs from the first argument in several ways. The first argument referred directly to the facts about Victor that show he died, whereas this second argument does not say which features of the victim show that it was Victor. Instead, this argument relies on trusting other people—Victor's friends—without knowing what it was about the victim's face that made them think it was Victor. Such appeals to authority will be discussed in more detail in Chapters 3 and 15.

The third issue is the cause of death. Here it is common to appeal to a medical authority. In our case, the coroner or medical examiner makes

observations or runs scientific tests that provide premises for another argument that is supposed to justify the conclusion that Victor's death was caused by a bullet to the head. This argument is also an appeal to an authority, but here the authority is a scientific expert rather than a friend.

Yet another argument, possibly based on firing marks on the bullet, can then justify you in believing that the bullet came from a certain gun. More arguments, possibly based on eyewitnesses, then justify the claims that Madison was the person who fired that gun at Victor. And so on.

All of these arguments depend on background assumptions. When you see the marks on the bullet that killed Victor line up with the marks on another bullet that was fired from the alleged murder weapon, you assume that guns leave distinctive marks on bullets and that nobody switched the bullets. A good prosecutor will provide arguments for these assumptions, but nobody can prove everything. Arguments always start from assumptions. This problem will occupy us at several points later, including parts of Chapters 3 and 5. The point for now is just that the prosecution needs to produce several arguments of various kinds in order to justify the claim that Madison killed Victor.

It is also crucial that killing violates the law. If not, then Madison should not be found guilty for killing Victor. So, how can the prosecutor justify the assumption that such killing is illegal? Prosecutors usually just quote a statute or cite a common law principle and apply it to the case, but that argument assumes a lot of background information. In the case of a statute, there must be a duly elected legislature, it must have jurisdiction over the place and time where and when the killing occurred, it must follow required procedures, and the content of the law must be constitutionally permissible. Given such a context, if the legislature says that a certain kind of killing is illegal, then it is illegal. It is fascinating that merely announcing that something is illegal thereby makes it illegal. We will explore such performatives and speech acts in Chapter 2. For now we will simply assume that all of these arguments could be provided if needed.

Even so, Madison might have had some justification for killing Victor, such as self-defense. This justification for her act can be presented in an argument basically like this: I have a reason to protect my own life, and I need to kill Victor first in order to protect my own life, so I have a reason to kill Victor. This justification differs in several ways from the kind of justification that we have been discussing so far. For one thing, this argument provides a reason for a different person—a reason for Madison—whereas the preceding arguments provided a reason for you as a juror. This argument also provides a reason with a different kind of object, since it justifies an action (killing Victor) whereas the previous arguments justified a belief (the belief that Madison did kill Victor). It provides a practical reason instead of an intellectual reason. Despite these differences, however, if her attorneys want to show that Madison has this new kind of justification, they need to give an argument to show that she was justified in doing what she did.

Even if Madison had no justification, she still might have had an excuse. Whereas a justification is supposed to show that the act was the right thing to do, an excuse admits that the act was wrong but tries to show that the agent was not fully responsible for doing it. Madison might, for example, argue that she honestly believed that Victor was going to kill her if she did not kill him first. If she offers this only as an excuse, she can admit that her belief was mistaken, so she had no justification for killing Victor. Her claim is, instead, that she was not fully responsible for his death because she was only trying to defend herself.

Excuses like this are, in effect, explanations. By citing her mistake, Madison explains why she did what she did. If she had killed Victor because she hated him or because she wanted to take his money, then she would have no excuse. Her act is less blameworthy, however, if she was mistaken. Of course, you should be careful before you shoot someone, so Madison could still be guilty of carelessness or negligence. But that is not as bad as killing someone out of hatred or for money. Her mistake might even be reasonable. If Victor was aiming a gun at her, then, even if it turned out not to be loaded, any rational person in her position might have thought that Victor was on the attack. Such reasonable mistakes might reduce or even remove responsibility. Thus, by explaining her act as a mistake, Madison puts her act in a better light than it would appear without that explanation. In general, an excuse is just an explanation of an act that puts that act in a better light by reducing the agent's responsibility.

To offer an excuse, then, Madison's defense attorneys will need to give arguments whose purpose is not justification but explanation. This excuse will then determine what she is guilty of. Whether Madison is guilty of first-degree murder or some lesser charge, such as second-degree murder or manslaughter, or even no crime at all, depends on the explanation for her act of killing Victor.

Several of the earlier arguments also provided explanations. The medical examiner cited the head wound to explain why Victor stopped breathing. The victim's identity explained why his friends said he was Victor. The fact that the bullet came out of a particular gun explained why it had certain markings. The legislature's vote explained why the killing was illegal. And so on.

In this way, what appears at first to be a simple case actually depends on a complex chain of arguments that mixes justifications with explanations. All of these justifications and explanations can be understood by presenting them explicitly in the form of arguments.

One final point is crucial. Suppose that Madison has no justification or excuse for killing Victor. It is still not enough for the prosecutor to give any old argument that Madison killed Victor. The prosecution must prove guilt beyond a reasonable doubt. This burden of proof makes the strength of the argument crucial. You as a juror should not convict, even if you think Madison is guilty, unless the prosecution's argument meets this high

standard. In this case, as in many others, it is not enough just to be able to identify the argument and to understand its purpose. You also need to determine how strong it is.

For such reasons, we all need to understand arguments and to be able to evaluate them. This need arises not only in law but also in life, such as when we decide which candidate to vote for, what course to take, whether to believe that your spouse is cheating on you, and so on. The goal of this book is to teach the skills needed for understanding and assessing arguments about important issues like these.

#### DISCUSSION QUESTION

In his famous testimony to the United Nations Security Council on February 5, 2003, which was forty-two days before U.S. troops entered Iraq, Secretary of State Colin Powell gave several arguments for his main conclusion that Saddam Hussein was at that time still trying to obtain fissile material for a nuclear weapons program. His arguments mix justification with explanation. For each of his arguments, determine whether it is a justification or an explanation. How does each argument work? How strong is it? How would you respond if you disagreed? How would you defend that part against criticisms? It will, of course, be difficult to answer these questions before studying the rest of this book. However, it is worthwhile to reflect on how much you already understand at the start. It is also useful to have some concrete examples to keep in mind as you study arguments in more depth.

Let me turn now to nuclear weapons. We have no indication that Saddam Hussein has ever abandoned his nuclear weapons program. On the contrary, we have more than a decade of proof that he remains determined to acquire nuclear weapons.

To fully appreciate the challenge that we face today, remember that in 1991 the inspectors searched Iraq's primary nuclear weapons facilities for the first time, and they found nothing to conclude that Iraq had a nuclear weapons program. But, based on defector information, in May of 1991, Saddam Hussein's lie was exposed. In truth, Saddam Hussein had a massive clandestine nuclear weapons program that covered several different techniques to enrich uranium, including electromagnetic isotope separation, gas centrifuge and gas diffusion.

We estimate that this illicit program cost the Iraqis several billion dollars. Nonetheless, Iraq continued to tell the IAEA that it had no nuclear weapons program. If Saddam had not been stopped, Iraq could have produced a nuclear bomb by 1993, years earlier than most worst case assessments that had been made before the war.

In 1995, as a result of another defector, we find out that, after his invasion of Kuwait, Saddam Hussein had initiated a crash program to build a crude

(continued)

nuclear weapon, in violation of Iraq's UN obligations. Saddam Hussein already possesses two out of the three key components needed to build a nuclear bomb. He has a cadre of nuclear scientists with the expertise, and he has a bomb design.

Since 1998, his efforts to reconstitute his nuclear program have been focused on acquiring the third and last component: sufficient fissile material to produce a nuclear explosion. To make the fissile material, he needs to develop an ability to enrich uranium. Saddam Hussein is determined to get his hands on a nuclear bomb.

He is so determined that he has made repeated covert attempts to acquire high-specification aluminum tubes from eleven different countries, even after inspections resumed. These tubes are controlled by the Nuclear Suppliers Group precisely because they can be used as centrifuges for enriching uranium.

By now, just about everyone has heard of these tubes and we all know that there are differences of opinion. There is controversy about what these tubes are for. Most U.S. experts think they are intended to serve as rotors in centrifuges used to enrich uranium. Other experts, and the Iraqis themselves, argue that they are really to produce the rocket bodies for a conventional weapon, a multiple rocket launcher.

Let me tell you what is not controversial about these tubes. First, all the experts who have analyzed the tubes in our possession agree that they can be adapted for centrifuge use.

Second, Iraq had no business buying them for any purpose. They are banned for Iraq.

I am no expert on centrifuge tubes, but this is an old army trooper. I can tell you a couple things.

First, it strikes me as quite odd that these tubes are manufactured to a tolerance that far exceeds U.S. requirements for comparable rockets. Maybe Iraqis just manufacture their conventional weapons to a higher standard than we do, but I don't think so.

Second, we actually have examined tubes from several different batches that were seized clandestinely before they reached Baghdad. What we notice in these different batches is a progression to higher and higher levels of specification, including in the latest batch an anodized coating on extremely smooth inner and outer surfaces.

Why would they continue refining the specifications? Why would they go to all that trouble for something that, if it was a rocket, would soon be blown into shrapnel when it went off?

The high-tolerance aluminum tubes are only part of the story. We also have intelligence from multiple sources that Iraq is attempting to acquire magnets and high-speed balancing machines. Both items can be used in a gas centrifuge program to enrich uranium.

In 1999 and 2000, Iraqi officials negotiated with firms in Romania, India, Russia and Slovenia for the purchase of a magnet production plant. Iraq wanted the plant to produce magnets weighing twenty to thirty grams. That's the same weight as the magnets used in Iraq's gas centrifuge program before the Gulf War.

This incident, linked with the tubes, is another indicator of Iraq's attempt to reconstitute its nuclear weapons program.

Intercepted communications from mid-2000 through last summer showed that Iraqi front companies sought to buy machines that can be used to balance gas centrifuge rotors. One of these companies also had been involved in a failed effort in 2001 to smuggle aluminum tubes into Iraq.

People will continue to debate this issue, but there is no doubt in my mind. These illicit procurement efforts show that Saddam Hussein is very much focused on putting in place the key missing piece from his nuclear weapons program, the ability to produce fissile material.



## THE WEB OF LANGUAGE

*Arguments are made up of language, so we cannot understand arguments without first understanding language. This chapter will examine some of the basic features of language, stressing three main ideas. First, language is conventional. Words acquire meaning within a rich system of linguistic conventions and rules. Second, the uses of language are diverse. We use language to communicate information, but we also use it to ask questions, issue orders, write poetry, keep score, formulate arguments, and perform an almost endless number of other tasks. Third, meaning is often conveyed indirectly. To understand the significance of many utterances, we must go beyond what is literally said to examine what is conversationally implied by saying it.*

### LANGUAGE AND CONVENTION

---

The preceding chapter stressed that arguing is a *practical* activity. More specifically, it is a *linguistic* activity. Arguing is one of the many things that we can do with words. In fact, unlike things that we can accomplish both with words and without words (like making people happy, angry, and so forth), arguing is something we can *only* do with words or other meaningful symbols. That is why nonhuman animals never give arguments. To understand how arguments work, then, it is crucial to understand how language works.

Unfortunately, our understanding of human language is far from complete, and linguistics is a young science in which disagreement exists on many important issues. Still, certain facts about language are beyond dispute, and recognizing them will provide a background for understanding how arguments work.

As anyone who has bothered to think about it knows, language is conventional. There is no reason why we, as English speakers, use the word “dog” to refer to a dog rather than to a cat, a tree, or the number of planets in our solar system. It seems that any word might have been used to stand for anything. Beyond this, there seems to be no reason why we put words together the way we do. In English, we put adjectives before the nouns they modify. We thus speak of a “green salad.” In French, adjectives usually follow the noun, and so, instead of saying “verte salade,” the French say “salade verte.”

The conventions of our own language are so much with us that it strikes us as odd when we discover that other languages have different conventions. A French diplomat once praised his own language because, as he said, it followed the natural order of thought. This strikes English speakers as silly, but in seeing why it is silly, we see that the word order in our own language is conventional as well.

Although it is important to realize that language is conventional, it is also important not to misunderstand this fact. From the idea that language is conventional, it is easy to conclude that language is totally arbitrary. If language is totally arbitrary, then it might seem that it really does not matter which words we use or how we put them together. It takes only a little thought to see that this view, however daring it might seem, misrepresents the role of conventions in language. If we wish to communicate with others, we must follow the system of conventions that others use. Grapefruits are more like big lemons than like grapes, so you might want to call them "mega-lemons." Still, if you order a glass of mega-lemon juice in a restaurant, you will get stares and smirks but no grapefruit juice. The same point lies behind this famous passage in *Through the Looking Glass*, by Lewis Carroll:

"There's glory for you!"

"I don't know what you mean by 'glory,'" Alice said.

Humpty Dumpty smiled contemptuously.

"Of course you don't—till I tell you. I meant 'there's a nice knock-down argument for you!'"

"But 'glory' doesn't mean 'a nice knock-down argument,'" Alice objected.

"When I use a word," Humpty Dumpty said, in a rather scornful tone, "it means just what I choose it to mean—neither more nor less."

"The question is," said Alice, "whether you can make words mean so many different things."

The point, of course, is that Humpty Dumpty cannot make a word mean whatever he wants it to mean, and he cannot communicate if he uses words in his own peculiar way without regard to what those words themselves mean. Communication can take place only within a shared system of conventions. Conventions do not destroy meaning by making it arbitrary; conventions bring meaning into existence.

A misunderstanding of the conventional nature of language can lead to pointless disputes. Sometimes, in the middle of a discussion, someone will declare that "the whole thing is just a matter of definition" or "what you say is true by your definition, false by mine." There are times when definitions are important and the truth of what is said turns on them, but usually this is not the case. Suppose someone has fallen off a cliff and is heading toward certain death on the rocks below. Of course, it is a matter of convention that we use the word "death" to describe the result of the sudden, sharp stop at the end of the fall. We might have used some other word—perhaps "birth"—instead. But it certainly will not help a person who is falling to his

certain death to shout out, "By 'birth' I mean death." It will not help even if *everyone* agrees to use these words in this new way. If we all decided to adopt this new convention, we would then say, "He is falling from the cliff to his certain birth" instead of "He is falling from the cliff to his certain death." But speaking in this way will not change the facts. It will not save him from perishing. It will not make those who care for him feel better.

The upshot of this simple example is that the *truth* of what we say is rarely just a matter of definition. Whether what we have said is true or not will depend, for the most part, on how things stand in the world. Abraham Lincoln, during his days as a trial lawyer, is reported to have cross-examined a witness like this:

"How many legs does a horse have?"

"Four," said the witness.

"Now, if we call a tail a leg, how many legs does a horse have?"

"Five," answered the witness.

"Nope," said Abe, "calling a tail a leg don't make it a leg."

In general, then, though the *meaning* of what we say is dependent on convention, the *truth* of what we say is not.

In the preceding sentence we used the qualifying phrase, "in general." To say that a claim holds *in general* indicates that there may be exceptions. This qualification is needed because sometimes the truth of what we say is simply a matter of definition. Take a simple example: The claim that a triangle has three sides is true by definition, because a triangle is defined as "a closed figure having three sides." Again, if someone says that sin is wrong, he or she has said something that is true by definition, for a sin is defined as, among other things, "something that is wrong." In unusual cases like these, things are true merely as a matter of convention. Still, in general, the truth of what we say is settled not by appealing to definitions but, instead, by looking at the facts. In this way, language is not arbitrary, even though it is conventional.

## LINGUISTIC ACTS

In the previous section we saw that a language is a system of shared conventions that allows us to communicate with one another. If we examine language, we will see that it contains many different kinds of conventions. These conventions govern what we will call linguistic acts, speech acts, and conversational acts. We will discuss linguistic acts first.

We have seen that words have meanings conventionally attached to them. The word "dog" is used conventionally to talk about dogs. Given what our words mean, it would be incorrect to call dogs "airplanes." Proper names are also conventionally assigned, for Harry Jones could have been named

Wilbur Jones. Still, given that his name is not Wilbur, it would be improper to call him Wilbur. Rules like these, which govern meaning and reference, can be called *semantic* rules.

Other conventions concern the ways words can be put together to form sentences. These are often called *syntactic* or *grammatical* rules. Using the three words "John," "hit," and "Harry," we can formulate sentences with very different meanings, such as "John hit Harry" and "Harry hit John." We recognize that these sentences have different meanings, because we understand the grammar of our language. This grammatical understanding also allows us to see that the sentence "Hit John Harry" has no determinate meaning, even though the individual words do. (Notice that "Hit John, Harry!" does mean something: It is a way of telling Harry to hit John.) Grammatical rules are important, for they play a part in giving a meaning to combinations of words, such as sentences.

Some of our grammatical rules play only a small role in this important task of giving meaning to combinations of words. It is bad grammar to say, "If I was you, I wouldn't do that," but it is still clear what information the person is trying to convey. What might be called stylistic rules of grammar are of relatively little importance for logic, but grammatical rules that affect the meaning or content of what is said are essential to logical analysis. Grammatical rules of this kind can determine whether we have said one thing rather than another, or perhaps failed to say anything at all and have merely spoken nonsense.

It is sometimes hard to tell what is nonsense. Consider "The horse raced past the barn fell." This sentence usually strikes people as nonsense when they hear it for the first time. To show them that it actually makes sense, all we need to do is insert two words: "The horse that was raced past the barn fell." Since English allows us to drop "that was," the original sentence means the same as the slightly expanded version. Sentences like these are called "garden path sentences," because the first few words "lead you down the garden path" by suggesting that some word plays a grammatical role that it really does not play. In this example, "The horse raced . . ." suggests at first that the main verb is "raced." That makes it hard to see that the main verb really is "fell."

Another famous example is "Buffalo buffalo buffalo." Again, this seems like nonsense at first, but then someone points out that "buffalo" can be a verb meaning "to confuse." The sentence "Buffalo buffalo buffalo" then means "North American bison confuse North American bison." Indeed, we can even make sense out of "Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo Buffalo buffalo." This means "North American bison from Buffalo, New York, that North American bison from Buffalo, New York, confuse also confuse North American bison from Buffalo, New York, that North American bison from Buffalo, New York, confuse."

Examples like these show that sentences can have linguistic meaning when they seem meaningless. To be meaningful, sentences need to follow both *semantic* conventions that govern meanings of individual words and also *syntactic* or *grammatical* conventions that lay down rules for combining words into meaningful wholes. When a sentence satisfies essential semantic and syntactic conventions, we will say that the person who uttered that sentence performed a *linguistic act*: The speaker said something meaningful in a language.<sup>1</sup> The ability to perform linguistic acts shows a command of a language. What the speaker says may be false, irrelevant, boring, and so on; but, if in saying it linguistic rules are not seriously violated, then that person can be credited with performing a linguistic act.

Later, in Chapters 13 and 14, we will look more closely at semantic and syntactic conventions, for they are common sources of fallacies and other confusions. In particular, we shall see how these conventions can generate fallacies of ambiguity and fallacies of vagueness. Before examining the defects of our language, however, we should first appreciate that language is a powerful and subtle tool that allows us to perform a wide variety of jobs important for living in the world.

#### EXERCISE I

Read each of the following sentences aloud. Did you perform a linguistic act? If so, explain what the sentence means and why it might not seem meaningful.

1. The old man the ship.
2. Colorless green ideas sleep furiously.
3. Time flies like an arrow. Fruit flies like bananas.
4. The cotton clothing is made of grows in Mississippi.
5. The square root of pine is tree.
6. The man who whistles tunes pianos.
7. “ ‘Twas brillig, and the slithy toves did gyre and gimble in the wabe.”  
(From Lewis Carroll)

And now some weird examples from Dan Wegner’s Hidden Brain Damage Scale. If these make sense to you, it might be a sign of hidden brain damage. If they don’t make sense, explain why:

8. People tell me one thing one day and out the other.
9. I feel as much like I did yesterday as I do today.
10. My throat is closer than it seems.
11. I’ve lost all sensation in my shirt.
12. There’s only one thing for me.

**DISCUSSION QUESTIONS**

- When someone hums (but does not sing) the “Star-Spangled Banner,” does she perform a linguistic act? Why or why not?
- Can a speaker mispronounce a word in a sentence without performing any linguistic act? Why or why not?

© Jeff Stahler/CSL, CartonStock Ltd.

**SPEECH ACTS**

When asked about the function of language, it is natural to reply that we use language to describe objects and communicate ideas. These are, however, only some of the purposes for which we use language. Other purposes become obvious as soon as we look at the ways in which our language actually works. Adding up a column of figures is a linguistic activity—though it is rarely looked at in this way—but it does not describe any objects (since numbers are not objects) or communicate any ideas to others. When I add the figures, I am not even communicating anything to myself; I am just trying to figure something out. A look at our everyday conversations produces a host of other examples of language being used for different purposes. Grammarians, for example, have divided sentences into various moods, among which are:

Indicative: Barry Bonds hit a home run.

Imperative: Get in there and hit a home run, Barry!

Interrogative: Did Barry Bonds hit a home run?

Expressive: Hurray for Barry Bonds!

The first sentence states a fact. We can use it to communicate information about something that Barry Bonds did. If we use it in this way, then what we say will be either true or false. Notice that none of the other sentences can be called either true or false even though they are all meaningful.

## PERFORMATIVES

The different types of sentences recognized by traditional grammarians show that we use language to do more than convey information, but they still give only a small sample of the wide variety of things that we can accomplish using language. Sometimes, for example, we use language to perform an action. In one familiar setting, if one person says, "I do," and another person says, "I do," and finally a third person says, "I now pronounce you husband and wife," the relationship between the first two people changes in a fundamental way: They thereby become married. With luck, they begin a life of wedded bliss, but they also alter their legal relationship. For example, they may now file joint income tax returns and may not legally marry other people without first getting divorced. The philosopher J. L. Austin labeled such utterances *performatives* in order to contrast performing an action with simply stating or describing something.<sup>2</sup>

Performatives come in a wide variety of forms. They are often in the first person (like "I do"), but not always. For example, "You're all invited to my house after the game" is in the second person, but uttering it performs the act of inviting. Even silence can amount to a performative act in special situations. When the chairperson of a meeting asks if there are any objections to a ruling and none is voiced, then the voters, through their silence, have accepted the ruling.

Because of this diversity of forms, it is not easy to formulate a definition that covers all performatives, so we will not even try to define performatives here. Instead, we will concentrate on one particularly clear subclass of performatives, which J. L. Austin called *explicit performatives*. All explicit performatives are utterances in the first-person singular indicative noncontinuous<sup>3</sup> present. But not all utterances of that form are explicit performatives. There is one more requirement:

An utterance of that form is an explicit performative if and only if it yields a true statement when plugged into the following pattern:

In saying "I \_\_\_\_" in appropriate circumstances, I thereby \_\_\_\_.

For example, "I congratulate you" expresses an explicit performative, because in saying "I congratulate you," I thereby congratulate you. Here a quoted expression occurs on the left side of the word "thereby" but not on the right side. This reflects the fact that the formula takes us from the words (which are quoted) to the world (the actual act that is performed). The *saying*, which is referred to on the left side of the pattern, amounts to the *doing* referred to on the right side of the word "thereby." We will call this the *thereby test* for explicit performatives.

The thereby test includes an important qualification: *The context of the utterance must be appropriate*. You have not congratulated anyone if you say, “I congratulate you,” when no one is around, unless you are congratulating yourself. Congratulations said by an actor in a play are not real congratulations, and so on.

Assuming an appropriate context, all of the following sentences meet the thereby test:

I promise to meet you tomorrow.

I bid sixty-six dollars. (Said at an auction)

I bid one club. (Said in a bridge game)

I resign from this club.

I apologize for being late.

Notice that it doesn’t make sense to *deny* any of these performatives. If someone says, “I bid sixty-six dollars,” it is not appropriate for someone to reply “No, you don’t” or “That’s false.” It could, however, be appropriate for someone to reply, “You can’t bid sixty-six dollars, because the bidding is already up to seventy dollars.” In this case, the person tried to make a bid, but failed to do so.

Several explicit performatives play important roles in constructing arguments. These include sentences of the following kind:

I *conclude* that this bill should be voted down.

I *base* my conclusion on the assumption that we do not want to hurt the poor.

I *stipulate* that anyone who earns less than \$10,000 is poor.

I *assure* you that this bill will hurt the poor.

I *concede* that I am not absolutely certain.

I *admit* that there is much to be said on both sides of this issue.

I *give my support* to the alternative measure.

I *deny* that this alternative will hurt the economy.

I *grant* for the sake of argument that some poor people are lazy.

I *reply* that most poor people contribute to the economy.

I *reserve comment* on other issues raised by this bill.

We will call this kind of performative an *argumentative performative*. Studying such argumentative performatives can help us to understand what is going on in arguments, which is one main reason why we are studying performatives here.

In contrast to the above utterances, which pass the thereby test, none of the following utterances does:

I agree with you. (This describes one’s thoughts or beliefs, so, unlike a performative, it can be false.)

I am sorry for being late. (This describes one's feelings and could be false.)

Yesterday I bid sixty dollars. (This is a statement about a past act and might be false.)

I'll meet you tomorrow. (This utterance may only be a prediction that can turn out to be false.)

Questions, imperatives, and exclamations are also not explicit performatives, because they cannot sensibly be plugged into the thereby test at all. They do not have the right form, since they are not in the first-person singular indicative noncontinuous present.

#### EXERCISE II

Using the thereby test as described above, indicate which of the following sentences express explicit performatives (EP) and which do not express explicit performatives (N) in appropriate circumstances:

1. I pledge allegiance to the flag.
2. We pledge allegiance to the flag.
3. I pledged allegiance to the flag.
4. I always pledge allegiance at the start of a game.
5. You pledge allegiance to the flag.
6. He pledges allegiance to the flag.
7. He doesn't pledge allegiance to the flag.
8. Pledge allegiance to the flag!
9. Why don't you pledge allegiance to the flag?
10. Pierre is the capital of South Dakota.
11. I state that Pierre is the capital of South Dakota.
12. I order you to leave.
13. Get out of here!
14. I didn't take it.
15. I swear that I didn't take it.
16. I won't talk to you.
17. I refuse to talk to you.
18. I'm out of gas.
19. I feel devastated.
20. Bummer!
21. I claim this land for England.
22. I bring you greetings from home.

## KINDS OF SPEECH ACTS

Recognizing explicit performatives introduces us to a kind of act distinct from linguistic acts. We will call them *speech acts*.<sup>4</sup> They include such acts as stating, promising, swearing, and refusing. A speech act is the conventional move that a remark makes in a language exchange. It is what is done *in* saying something.

Speech acts are distinct from linguistic acts, because the same linguistic act can play different roles in different contexts. This is shown by the following brief conversations.

A: Is there any pizza left?

B: Yes.

A: Do you promise to pay me back by Friday?

B: Yes.

A: Do you swear to tell the truth?

B: Yes.

A: Do you refuse to leave?

B: Yes.

Here the same linguistic act, uttering the word “yes,” is used to do four different things: to state something, to make a promise, to take an oath, and to refuse to do something.

We can make this idea of a speech act clearer by using the notion of an explicit performative. The basic idea is that different speech acts are named by the different verbs that occur in explicit performatives. We can thus use the thereby test to search for different kinds of speech acts. For example:

If I say, “I promise,” I thereby promise. So “I promise” is a performative, and *promising* is a kind of speech act.

If I say, “I resign,” I thereby resign. So “I resign” is a performative, and *resigning* is a kind of speech act.

If I say, “I apologize,” I thereby apologize. So “I apologize” is a performative, and *apologizing* is a kind of speech act.

If I say, “I question his honesty,” I thereby question his honesty. So “I question his honesty” is a performative, and *questioning* is a kind of speech act.

If I say, “I conclude that she is guilty,” I thereby conclude that she is guilty. So “I conclude that she is guilty” is a performative, and *concluding* is a kind of speech act.

The main verbs that appear in such explicit performatives can be called *performative verbs*. Performative verbs name kinds of speech acts.

Still, the same speech act can also be accomplished without any performative verb. I can deny my opponent's claim by saying either "I deny that" or simply "No way!" Both utterances perform the speech act of denying, even though only the former is a performative. The latter is not a performative and does not contain any performative verb, but it still performs a speech act.

Nonetheless, speech acts depend on context in much the same way as performatives. If a baseball umpire during a game shouts, "You're out!" to a batter then the batter is out. By way of contrast, if someone in the stands shouts, "You're out!" or "He's out!" the batter is not *thereby* out, although the person who shouts this may be encouraging the umpire to call the batter out or complaining because he didn't. And even an umpire cannot call a player out if the player is not at bat, but is pitching or in the dugout. The identity of the speaker and the audience as well as the circumstances thus determines whether the speech act is accomplished. Similarly, in a less formal setting, I cannot invite someone to your party (unless you gave me permission to do so), and I cannot congratulate you for losing your job (at least not sincerely). This example shows that a speech act will *fail to come off* or will be *void* unless certain rules or conventions are satisfied. These rules or conventions that must be satisfied for a speech act to come off and not be void can be called *speech act rules*.

### EXERCISE III

Which of the following verbs names a speech act?

1. *capture* the suspect
2. *assert* that the suspect is guilty
3. *stare* accusingly at the suspect
4. *find* the defendant guilty
5. *punish* the defendant
6. *take* the defendant away
7. *revoke* the defendant's driver's license
8. *welcome* the prisoner to prison
9. *order* the prisoner to be silent
10. *lock* the cell door

### EXERCISE IV

Using a dictionary, find ten verbs that can be used to construct explicit performatives that have not yet been mentioned in this chapter.

**DISCUSSION QUESTIONS**

1. Do the speech acts in which people get married presuppose that the people who are getting married are of different sexes? *Should* these speech acts presuppose this fact? Why or why not?
2. The importance of deciding what kind of speech act has been performed is illustrated by a classic case from the law of contracts, *Hawkins v. McGee*.<sup>5</sup> McGee performed an operation on Hawkins that proved unsuccessful, and Hawkins sued for damages. He did not sue on the basis of malpractice, however, but on the basis of breach of contract. His attorney argued that the doctor initiated a contractual relationship in that he tried to persuade Hawkins to have the operation by saying things such as "I will guarantee to make the hand a hundred percent perfect hand." He made statements of this kind a number of times, and Hawkins finally agreed to undergo the operation on the basis of these remarks. Hawkins's attorney maintained that these exchanges, which took place in the doctor's office on a number of occasions, constituted an offer of a contract that Hawkins explicitly accepted. The attorney for the surgeon replied that these words, even if uttered, would not constitute an offer of a contract, but merely expressed a *strong belief*, and that reasonable people should know that doctors cannot guarantee results.

It is important to remember that contracts do not have to be written and signed to be binding. A proper verbal offer and acceptance are usually sufficient to constitute a contract. The case, then, turned on two questions: (1) Did McGee utter the words attributed to him? In other words, did McGee perform the *linguistic act* attributed to him? The jury decided that he did. (2) The second, more interesting question was whether these words, when uttered in this particular context, amounted to an offer of a contract, as Hawkins's attorney maintained, or merely were an expression of strong belief, as McGee's attorney held. In other words, the fundamental question in this case was what kind of *speech act* McGee performed when trying to convince Hawkins to have the operation.

Explain how you would settle this case. (The court actually ruled in favor of Hawkins, but you are free to disagree.)

**CONVERSATIONAL ACTS**

In examining linguistic acts (saying something meaningful in a language) and then speech acts (doing something in using words), we have largely ignored a central feature of language: It is normally a *practical* activity with certain goals. We use language in order to inform people of things, get them to do things, amuse them, calm them down, and so on. We can capture this practical aspect of language by introducing the notion of a *conversational exchange*, that is, a situation where various speakers use speech acts in order to bring about some effects in each other. We will call this act of using a speech act to cause a standard effect in another a *conversational act*.

Suppose, for example, Amy says to Bobbi, "Someone is following us." In this case, Amy has performed a linguistic act; that is, she has uttered a meaningful sentence in the English language. Amy has also performed a speech act—specifically, she has *stated* that they are being followed. The point of performing this speech act is to produce in Bobbi a particular belief—namely, that they are being followed. (Amy's utterance might also have other purposes, such as to alert Bobbi to some danger, but it accomplishes those other purposes by means of getting Bobbi to believe they are being followed.) If Amy is successful in this, then Amy has successfully performed the conversational act of producing this belief in Bobbi. Amy, of course, might fail in her attempt to do this. Amy's linguistic act could be successful and her speech act successful as well, yet, for whatever reason, Bobbi might not accept as true what Amy is telling her. Perhaps Bobbi thinks that Amy is paranoid or just trying to frighten her as some kind of joke. In that case, Amy failed to perform her intended conversational act, even though she did perform her intended linguistic and speech acts.

Here are some other examples of the difference between performing a speech act and performing a conversational act:

We can *warn* people about something in order to *put them on guard* concerning it.

Here warning is the speech act; putting them on guard is the intended conversational act.

We can *urge* people to do things in order to *persuade* them to do these things. Here urging is the speech act; persuading is the intended conversational act.

We can *assure* people concerning something in order to *instill confidence in them*.

Here assuring is the speech act; instilling confidence is the intended conversational act.

We can *apologize* to people in order to *make them feel better about us*.

Here apologizing is the speech act; making them feel better about us is the intended conversational act.

In each of these cases, our speech act may not succeed in having its intended conversational effect. Our urging, warning, and assuring may, respectively, fail to persuade, put on guard, or instill confidence. Indeed, speech acts may bring about the opposite of what was intended. People who brag (a speech act) in order to impress others (the intended conversational act) often actually make others think less of them (the actual effect). In many ways like these, we can perform a speech act without performing the intended conversational act.

The relationship between conversational acts and speech acts is confusing, because both of them can be performed at once by the same utterance. Suppose Carl says, "You are invited to my party." By means of this single

utterance, he performs a linguistic act of uttering this meaningful sentence, a speech act of inviting you, and perhaps also a conversational act of getting you to come to his party. Indeed, he would not be able to perform this conversational act without also performing such a speech act, assuming that you would not come to his party if you were not invited. He would also not be able to perform this speech act without performing this linguistic act or something like it, since he cannot invite you by means of an inarticulate grunt or by asking, "Are you invited to my party?"

As a result, we cannot sensibly ask whether Carl's utterance of "You are invited to my party" is a linguistic act, a speech act, or a conversational act. That single utterance performs all three acts at once. Nonetheless, we can distinguish those kinds of acts that Carl performs in terms of the verbs that describe the acts. Some verbs describe speech acts; other verbs describe conversational acts. We can tell which verbs describe which kinds of acts by asking whether the verb passes the thereby test (in which case the verb describes a speech act) or whether, instead, it describes a standard effect of the utterance (in which case the verb describes a conversational act).

#### EXERCISE V

Indicate whether the verbs in the following sentences name a speech act, a conversational act, or neither. Assume a standard context. Explain your answers.

1. She *thought* that he did it.
2. She *asserted* that he did it.
3. She *convinced* them that he did it.
4. She *condemned* him in front of everyone.
5. She *challenged* his integrity.
6. She *embarrassed* him in front of them.
7. He *denied* doing it.
8. They *believed* her.
9. They *encouraged* him to admit it.
10. She *told* him to get lost.
11. He *praised* her lavishly.
12. His *praise* *made* her happy.
13. He *threatened* to reveal her secret.
14. He *submitted* his resignation.
15. Her news *frightened* him half to death.
16. He *advised* her to go into another line of work.
17. She *blamed* him for her troubles.
18. His lecture *enlightened* her.
19. His jokes *amused* her.
20. His book *confused* her.

## CONVERSATIONAL RULES

Just as there are rules that govern linguistic acts and other rules that govern speech acts, so too there are rules that govern conversational acts. This should not be surprising, because conversations can be complicated interpersonal activities in need of rules to make them effective in attaining their goals. These underlying rules are implicitly understood by users of the language, but the philosopher Paul Grice was the first person to examine them in careful detail.<sup>6</sup>

We can start by examining standard or normal conversational exchanges where conversation is a cooperative venture—that is, where the people involved in the conversation have some common goal they are trying to achieve in talking with one another. (A prisoner being interrogated and a shop owner being robbed are *not* in such cooperative situations.) According to Grice, such exchanges are governed by what he calls the *Cooperative Principle*. This principle states that the parties involved should use language in a way that contributes toward achieving their common goal. It tells them to cooperate.

This general principle gains more content when we consider other forms of cooperation. Carpenters who want to build a house need enough nails and wood, but not too much. They need the right kinds of nails and wood. They also need to put the nails and wood together in the relevant way—that is, according to their plans. And, of course, they also want to perform their tasks quickly and in the right order. Rational people who want to achieve common goals must follow similar general restrictions in other practical activities. Because cooperative conversations are one such practical activity, speakers who want to cooperate with one another must follow rules analogous to those for carpenters.

Grice spells out four such rules. The first he calls the rule of *Quantity*. It tells us to give the right amount of information. More specifically:

1. Make your contribution as informative as is required (for the current purposes of the exchange);

and possibly:

2. Do not make your contribution more informative than is required.

Here is an application of this rule: A person rushes up to you and asks, "Where is a fire extinguisher?" You know that there is a fire extinguisher five floors away in the basement, and you also know that there is a fire extinguisher just down the hall. Suppose you say that there is a fire extinguisher in the basement. Here you have said something true, but you have violated the first part of the rule of *Quantity*. You have failed to reveal an important piece of information that, under the rule of *Quantity*, you should have produced. A violation of the second version of the rule would look like this: As smoke billows down the hall, you say where a fire extinguisher is located on each floor, starting with the basement. Eventually you will get around to

saying that there is a fire extinguisher just down the hall, but you bury the point in a mass of unnecessary information.

Grice's second rule is called the rule of *Quality*. In general: Try to make your contribution one that is true. More specifically:

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

In a cooperative activity, you are not supposed to tell lies. Beyond this, you are expected not to talk off the top of your head either. When we make a statement, we can be challenged by someone asking, "Do you really believe that?" or "Why do you believe that?" That a person has the right to ask such questions shows that statement making is governed by the rule of Quality.

In a court of law, witnesses promise to tell the whole truth and nothing but the truth. The demand for *nothing but the truth* reflects the rule of Quality. The demand for *the whole truth* roughly reflects the rule of Quantity. Obviously, nobody really tells every truth he or she knows. Here the *whole truth* concerns all the known truths that are relevant in the context.

This brings us to our next rule, the rule of *Relevance*. Simply stated, the rule of Relevance says:

Be relevant!

Though easy to state, the rule is not easy to explain, because relevance itself is a difficult notion. It is, however, easy to illustrate. If someone asks me where he can find a doctor, I might reply that there is a hospital on the next block. Though not a direct answer to the question, it does not violate the rule of Relevance because it provides a piece of useful information. If, however, in response I tell the person that I like his haircut, then I have violated the rule of Relevance. Clear-cut violations of this principle often involve *changing the subject*.

Another rule concerns the manner of our conversation. We are expected to be clear in what we say. Under the general rule of *Manner* come various special rules:

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief.
4. Be orderly.

As an example of the fourth part of this rule, when describing a series of events, it is usually important to state them in the order in which they occurred. It would certainly be misleading to say that two people had a child and got married when, in fact, they had a child after they were married.

Many other rules govern our conversations. "Be polite!" is one of them. "Be charitable!" is another. That is, we should put the best interpretation on what others say, and our replies should reflect this. We should avoid quibbling and being picky. For the most part, however, we will not worry about these other rules.

## EXERCISE VI

Indicate which, if any, of Grice's conversational rules are violated by the italicized sentence of each of the following conversations. Assume a standard context. More than one rule might be violated.

1. "Did you like her singing?" "*Her costume was beautiful.*"
2. "*The governor has the brains of a three-year-old.*"
3. "*The Lone Ranger rode into the sunset and jumped on his horse.*"
4. "*Without her help, we'd be up a creek without a paddle.*"
5. "Where is Palo Alto?" "*On the surface of the Earth.*"
6. "*It will rain tomorrow.*" "How do you know?" "I just guessed."
7. "*Does the dog need to go out for a W-A-L-K [spelled out]?*"
8. "Why did the chicken cross the road?" "*To get to the other side.*"



"What are you implying?"

## CONVERSATIONAL IMPLICATION

In a normal setting where people are cooperating toward reaching a shared goal, they often conform quite closely to Grice's conversational rules. If, on the whole, people did not do this, we could not have the linguistic practices we do. If we thought, for example, that people very often lied (even about the most trivial matters), the business of exchanging information would be badly damaged.

Still, people do not always follow these conversational rules. They withhold information, they elaborate needlessly, they assert what they know to be false, they say the first thing that pops into their heads, they wander off the subject, and they talk vaguely and obscurely. When we observe actual conversations, it is sometimes hard to tell how any information gets communicated at all.

The explanation lies in the same conversational rules. Not only do we usually follow these conventions, we also (1) implicitly realize that we are following them, and (2) expect others to assume that we are following them. This mutual understanding of the commitments involved in a conversational act has the following important consequence: People are able to convey a great deal of information without actually saying it.

A simple example will illustrate this point. Again suppose that a person, with smoke billowing behind him, comes running up to you and asks, "Where's a fire extinguisher?" You reply, "There's one in the lobby." Through a combination of conversational rules, notably relevance, quantity, and manner, this commits you to the claim that this is the closest, or at least the most accessible, fire extinguisher. Furthermore, the person you are speaking to assumes that you are committed to this. Of course, you have not actually *said* that it is the closest fire extinguisher; but you have, we might say, *implied* this. When we do not actually say something but imply it by virtue of a mutually understood conversational rule, the implication is called a *conversational implication*.

It is important to realize that conversational implication is a pervasive feature of human communication. It is not something we employ only occasionally for special effect. In fact, virtually every conversation relies on these implications, and most conversations would fall apart if people refused to go beyond literal meanings to take into account the implications of saying things. In the following conversation, *B* is literal-minded in just this way:

- A:** Do you know what time it is?  
**B:** Not without looking at my watch.

*B* has answered *A*'s question, but it is hard to imagine that *A* has received the information she was looking for. Presumably, she wanted to know what time it was, not merely whether *B*, at that very moment, knew the time. Finding *B* rather obtuse, *A* tries again:

- A:** Can you tell me what time it is?  
**B:** Oh, yes, all I have to do is look at my watch.

Undaunted, *A* gives it another try:

- A:** Will you tell me what time it is?  
**B:** I suppose I will as soon as you ask me.

Finally:

- A:** What time is it?  
**B:** Two o'clock. Why didn't you ask me that in the first place?

Notice that in each of these exchanges *B* gives a direct and accurate answer to *A*'s question; yet, in all but the last answer, *B* does not provide *A* with what *A* wants. Like a computer in a science-fiction movie, *B* is taking *A*'s questions too literally. More precisely, *B* does nothing *more* than take *A*'s remarks literally. In a conversational exchange, we expect others to take our remarks in the light of the obvious purpose we have in making them. We expect them to share our commonsense understanding of why people ask questions. At the very least, we expect people to respond to us in ways that are *relevant* to our purposes. Except at the end, *B* seems totally oblivious to the point of *A*'s questions. That is what makes *B* unhelpful and annoying.

Though all the conversational rules we have examined can be the basis of conversational implication, the rule of Relevance is particularly powerful in this respect. Normal conversations are dense with conversational implications that depend on the rule of Relevance. Someone says, "Dinner's ready," and that is immediately taken to be a way of asking people to come to the table. Why? Because dinner's being ready is a transparent *reason* to come to the table to eat. This is an ordinary context that most people are familiar with. Change the context, however, and the conversational implications can be entirely different. Suppose the same words, "Dinner's ready," are uttered when guests have failed to arrive on time. In this context, the conversational implication, which will probably be reflected in an annoyed tone of voice, will be quite different.

#### EXERCISE VII

Assuming a natural conversational setting, what might a person intend to conversationally imply by making the following remarks? Briefly explain why each of these conversational implications holds; that is, explain the relationship between what the speaker *literally* says and what the speaker intends to convey through conversational implication. Finally, for each example, find a context where the standard conversational implication would fail and another arise in its place.

1. It's getting a little chilly in here. (Said by a visitor in your home)
2. Do you mind if I borrow your pen? (Said to a friend while studying)
3. We are out of soda. (Said by a child to her parents)
4. I got here before he did. (Said in a ticket line)
5. Don't blame me if you get in trouble. (Said by someone who advised you not to do what you did)
6. Has this seat been taken? (Said in a theater before a show)
7. Don't ask me. (Said in response to a question)
8. I will be out of town that day. (Said in response to a party invitation)



GARFIELD © 1999 Paws, Inc. Reprinted with permission of UNIVERSAL UCLICK All rights reserved.

## RHETORICAL DEVICES

Many rhetorical devices work by openly violating conversational rules in order to generate conversational implications. Consider exaggeration. When someone claims to be hungry enough to eat a horse, it does not dawn on us to treat this as a literal claim about how much she can eat. To do so would be to attribute to the speaker a blatant violation of Grice's first rule of Quality—namely, do not say what you believe to be false. Consequently, her audience will naturally interpret her remark figuratively, rather than literally. They will assume that she is exaggerating the amount she can eat in order to conversationally imply that she is very hungry. This rhetorical device is called *overstatement* or *hyperbole*. It is commonly employed, often in heavy-handed ways.

Sometimes, then, we do not intend to have others take our words at face value. Even beyond this, we sometimes expect our listeners to interpret us as claiming just the *opposite* of what we assert. This occurs, for example, with *irony* and *sarcasm*. Suppose at a crucial point in a game, the second baseman fires the ball ten feet over the first baseman's head, and someone shouts, "Great throw." Literally, it was not a great throw; it was the opposite of a great throw, and this is just what the person who says "Great throw" is indicating. How do the listeners know they are supposed to interpret it in this way? Sometimes this is indicated by tone of voice. A sarcastic tone of voice usually indicates that the person means the opposite of what he or she is saying. Even without the tone of sarcasm, the remark "Great throw" is not likely to be taken literally. The person who shouts this knows that it was not a great throw, as do the people who hear it. Rather than attributing an obviously false belief to the shouter, we assume that the person is blatantly violating the rule of Quality to draw our attention to just how bad the throw really was.

Metaphors and similes are perhaps the most common forms of figurative language. A *simile* is, roughly, an explicit figurative comparison. A word such as "like" or "as" makes the comparison explicit, and the comparison is figurative because it would be inappropriate if taken literally. To say that the home team fought like tigers does not mean that they clawed the opposing team and took large bites out of them. To call someone as dumb as a post is not to claim that they have no brain at all.

With a *metaphor*, we also compare certain items, but without words such as "like" or "as." Metaphorical comparisons are still figurative because the vocabulary, at a literal level, is not appropriate to the subject matter. George Washington was not literally the father of his country. Taken literally, it hardly makes sense to speak of someone fathering a country. But the metaphor is so natural (or so familiar) that it does not cross our minds to treat the remark literally, asking, perhaps, who the mother was.

Taken literally, metaphors are usually obviously false, and then they violate Grice's rule of Quality. Again, as with irony, when someone says something obviously false, we have to decide what to make of that person's utterance. Perhaps the person is very stupid or a very bad liar, but often neither suggestion is plausible. In such a situation, sometimes the best supposition is that the person is speaking metaphorically rather than literally.

#### EXERCISE VIII

Identify each of the following sentences as irony, metaphor, or simile. For each sentence, write another expressing its literal meaning.

1. He missed the ball by a mile.
2. He acted like a bull in a china shop.
3. The exam blew me away.
4. He had to eat his words.
5. It was a real team effort. (Said by a coach after his team loses by forty points)
6. They are throwing the baby out with the bathwater.
7. This is a case of the tail wagging the dog.
8. "Religion is the opiate of the masses." (Marx)

#### EXERCISE IX

Unpack the following political metaphors by giving their literal content:

1. We can't afford a president who needs on-the-job training.
2. It's time for people on the welfare wagon to get off and help pull.
3. If you can't stand the heat, get out of the kitchen.
4. We need to restore a level playing field.
5. The special interests have him in their pockets.
6. He's a lame duck.

**DISCUSSION QUESTION**

At the start of the U.S. war with Iraq in 2003, some described Iraq as another Vietnam, while others described Saddam Hussein (Iraq's president) as another Hitler. Which metaphor was used by supporters of the war? Which was used by opponents? How can you tell? How do these metaphors work?

**SUMMARY**

In this chapter we have developed a rather complex picture of the way our language functions. In the process, we have distinguished three kinds or levels of acts that are performed when we employ language. We have also examined the rules associated with each kind or level of act. The following table summarizes this discussion:

**THREE LEVELS OF LANGUAGE**

<b>Kinds of Acts</b>	<b>Governing Rules</b>
A LINGUISTIC ACT is an act of saying something meaningful in a language. It is the basic act that is needed to make anything part of language.	Semantic rules (such as definitions) and syntactic rules (as in grammar).
A SPEECH ACT concerns the move a person makes in saying something. Different kinds of speech acts are indicated by the various verbs found in explicit performatives.	Speech act rules about special agents and circumstances appropriate to different kinds of speech acts.
A CONVERSATIONAL ACT is a speaker's act of causing a standard kind of effect in the listener; it is what I do by saying something—for example, I persuade someone to do something.	Conversational rules (the Cooperative Principle; Quantity, Quality, Relevance, and Manner).

**EXERCISE X**

1. It is late, and *A* is very hungry. *A* asks *B*, "When will dinner be ready?" Describe the linguistic act, the speech act, and some of the conversational acts this person may be performing in this context.
2. Someone is trying to solve the following puzzle: One of thirteen balls is heavier than the others, which are of equal weight. In no more than three weighings on a balance scale, determine which ball is the heavier one. The person is stumped, so someone says to her: "Begin by putting four balls in each pan of the scale." Describe the linguistic act, the speech act, and the conversational act of the person who makes this suggestion.

## NOTES

<sup>1</sup>J. L. Austin used the phrase “locutionary act” to refer to a level of language closely related to what we refer to as a “linguistic act.” See J. L. Austin, *How to Do Things with Words*, 2nd ed. (Cambridge, MA: Harvard University Press, 1975), 94–109.

<sup>2</sup>See, for example, J. L. Austin’s *How to Do Things with Words*.

<sup>3</sup>An example of the continuous present is “I bet ten dollars every week in the lottery.” Since this sentence is not used to make a bet, this sentence and others with the continuous present do not pass the thereby test or express explicit performatives.

<sup>4</sup>Austin calls speech acts “illocutionary acts.” See *How to Do Things with Words*, 98–132.

<sup>5</sup>Supreme Court of New Hampshire, 1929, 84 N.H. 114, A. 641.

<sup>6</sup>This discussion of conversational rules and implications is based on Paul Grice’s important essay, “Logic and Conversation,” which appears as the second chapter of his *Studies in the Way of Words* (Cambridge, MA: Harvard University Press, 1989). To avoid British references that some readers might find perplexing, we have sometimes altered Grice’s wording.



## THE LANGUAGE OF ARGUMENT

*Using the techniques developed in Chapter 2, this chapter will examine the use of language to formulate arguments and will provide methods to analyze genuine arguments in their richness and complexity. The first stage in analyzing an argument is the discovery of its basic structure. To do this, we will examine the words, phrases, and special constructions that indicate the premises and conclusions of an argument. The second stage is the study of techniques used to protect an argument. These include guarding premises so that they are less subject to criticism, offering assurances concerning debatable claims, and discounting possible criticisms in advance.*

### ARGUMENT MARKERS

---

In Chapter 2, we saw that language is used for a great many different purposes. One important thing that we do with language is construct arguments. Arguments are constructed out of statements, but arguments are not just lists of statements. Here is a simple list of statements:

Socrates is a man.

All men are mortal.

Socrates is mortal.

This list is not an argument, because none of these statements is presented as a reason for any other statement. It is, however, simple to turn this list into an argument. All we have to do is to add the single word “therefore”:

Socrates is a man.

All men are mortal.

Therefore, Socrates is mortal.

Now we have an argument. The word “therefore” converts these sentences into an argument by signaling that the statement following it is a *conclusion*, and the statement or statements that come before it are offered as *reasons* on behalf of this conclusion. The argument we have produced in this way is a good one, because the conclusion follows from the reasons stated on its behalf.

There are other ways of linking these sentences to form an argument. Here is one:

Since Socrates is a man,  
and all men are mortal,  
Socrates is mortal.

Notice that the word “since” works in roughly the opposite way that “therefore” does. The word “therefore” is a *conclusion marker*, because it indicates that the statement that follows it is a conclusion. In contrast, the word “since” is a *reason marker*, because it indicates that the following statement or statements are reasons. In our example, the conclusion comes at the end, but there is a variation on this. Sometimes the conclusion is given at the start:

Socrates is mortal, since all men are mortal and Socrates is a man.

“Since” flags reasons; the remaining connected statement is then taken to be the conclusion, whether it appears at the beginning or at the end of the sentence.

Many other terms are used to introduce an argumentative structure into language by marking either reasons or conclusions. Here is a partial list:

REASON MARKERS	CONCLUSION MARKERS
since	therefore
because	hence
for	thus
as	then

We shall call such terms “argument markers,” because each presents one or more statements as part of an argument or backing for some other statement.

It is important to realize that these words are not always used as argument markers. The words “since” and “then” are often used as indicators of time, as in, “He’s been an American citizen since 1973” and “He ate a hot dog, then a hamburger.” The word “for” is often used as a preposition, as in “John works for IBM.” Because some of these terms have a variety of meanings, it is not possible to identify argument markers in a mechanical way just by looking at words. It is necessary to examine the function of words in the context in which they occur. One test of whether a word is functioning as an argument marker in a particular sentence is whether you can substitute another argument marker without changing the meaning of the sentence. In the last example, it makes no sense to say, “John works since IBM.”

Many *phrases* are also available to signal that an argument is being given. Here is just a small sample:

- from which it follows that . . .
- from which we may conclude that . . .
- from which we see that . . .
- which goes to show that . . .
- which establishes that . . .

We can also indicate conclusions and reasons by using *argumentative performatives*, which we examined briefly in Chapter 2. If someone says, “I conclude that . . .,” the words that follow are given the status of a conclusion. More pretentiously, if someone says, “Here I base my argument on the claim that . . .,” what comes next has the status of a reason.

Examination of actual arguments will show that we have a great many ways of introducing an argumentative structure into our language by using the two forms of argument markers: reason markers and conclusion markers. The first, and in many ways the most important, step in analyzing an argument is to identify the conclusion and the reasons given on its behalf. We do this by paying close attention to these argument markers.

### **IF . . . , THEN . . .**

If-then sentences, which are also called *conditionals*, often occur in arguments, but they do not present arguments by themselves. To see this, consider the following conditional:

If the Dodgers improve their hitting, then they will win the Western Division.

The sentence between the “if” and the “then” is called the *antecedent* of the conditional. The sentence after the “then” is called its *consequent*. In uttering such a conditional, we are not asserting the truth of its antecedent, and we are not asserting the truth of its consequent either. Thus, the person who makes the above remark is not claiming that the Dodgers will win the Western Division. All she is saying is that *if* they improve their hitting, *then* they will win. Furthermore, she is not saying that they will improve their hitting. Because the speaker is not committing herself to either of these claims, she is not presenting an argument. This becomes clear when we contrast this conditional with a statement that does formulate an argument:

**CONDITIONAL:** *If* the Dodgers improve their hitting, *then* they will win the Western Division.

**ARGUMENT:** *Since* the Dodgers will improve their hitting, they will win the Western Division.

The sentence that follows the word “since” is asserted. That is why “since” is an argument marker, whereas the connective “if . . . then . . .” is not an argument marker.

Even though conditionals by themselves do not mark arguments, there is a close relationship between conditionals and arguments: Indicative conditionals provide *patterns* that can be converted into an argument whenever the antecedent is said to be true. (We also get an argument when the consequent is said to be false, but we will focus here on the simpler case of asserting the antecedent.) Thus, we often hear people argue in the following way:

If inflation continues to grow, there will be an economic crisis. But inflation will certainly continue to grow, so an economic crisis is on the way.

The first sentence is an indicative conditional. It makes no claims one way or the other about whether inflation will grow or whether an economic crisis will occur. The next sentence asserts the antecedent of this conditional and then draws a conclusion signaled by the argument marker “so.” We might say that when the antecedent of an indicative conditional is found to be true, the conditional can be *cashed in* for an argument.

Often the antecedent of a conditional is not asserted explicitly but is conversationally implied. When asked which player should be recruited for a team, the coach might just say, “If Deon is as good as our scouts say he is, then we ought to go for Deon.” This conditional does not actually assert that Deon is as good as the scouts report. Nonetheless, it would be irrelevant and pointless for the coach to utter this conditional alone if he thought that the scouts were way off the mark. The coach might immediately add that he disagrees with the scouting reports. But unless the coach cancels the conversational implication in some way, it is natural to interpret him as giving an argument that we ought to pick Deon. In such circumstances, then, an indicative conditional can conversationally imply an argument, even though it does not state the argument explicitly.

This makes it easy to see why indicative conditionals are a useful feature of our language. By providing patterns for arguments, they prepare us to draw conclusions when the circumstances are right. Much of our knowledge of the world around us is contained in such conditionals. Here is an example: If your computer does not start, the plug might be loose. This is a useful piece of practical information, for when your computer does not start, you can immediately infer that the plug might be loose, so you know to check it out.

Other words function in similar ways. When your computer fails to start, a friend might say, “Either the plug is loose or you are in deep trouble.” Now, if you also assert, “The plug is not loose,” you can conclude that you are in deep trouble. “Either . . . or . . .” sentences thus provide patterns for arguments, just as conditionals do. However, neither if-then sentences nor either-or sentences by themselves explicitly assert enough to present a complete argument, so “if . . . , then . . .” and “either . . . or . . .” should not be labeled as argument markers.

**EXERCISE I**

Indicate which of the following italicized words or phrases is a reason marker, a conclusion marker, or neither.

1. He apologized, so you should forgive him.
2. He apologized. *Accordingly*, you should forgive him.
3. *Since* he apologized, you should forgive him.
4. *Provided that* he apologized, you should forgive him.
5. *In view of the fact that* he apologized, you should forgive him.
6. He apologized. *Ergo*, you should forgive him.
7. *Given that* he apologized, you should forgive him.
8. He apologized, and *because of that* you should forgive him.
9. *After* he apologizes, you should forgive him.
10. He apologized. *As a result*, you should forgive him.
11. *Seeing as* he apologized, you should forgive him.
12. He apologized. *For that reason alone*, you should forgive him.

**EXERCISE II**

Indicate whether each of the following sentences is an argument.

1. Charles went bald, and most men go bald.
2. Charles went bald because most men go bald.
3. My roommate likes to ski, so I do, too.
4. My roommate likes to ski, and so do I.
5. I have been busy since Tuesday.
6. I am busy, since my teacher assigned lots of homework.

---

**ARGUMENTS IN STANDARD FORM**

---

Because arguments come in all shapes and forms, it will help to have a standard way of presenting arguments. For centuries, logicians have used a format of the following kind:

- (1) All men are mortal.
- (2) Socrates is a man.  
\_\_\_\_\_  
∴ (3) Socrates is mortal. (from 1–2)

The reasons (or premises) are listed and numbered. Then a line is drawn below the premises. Next, the conclusion is numbered and written below the line. The symbol “∴”—which is read “therefore”—is then added to the left of the conclusion in order to indicate the relation between the premises and the conclusion. Finally, the premises from which the conclusion is supposed to be derived are indicated in parentheses. Arguments presented in this way are said to be in *standard form*.

The notion of a standard form is useful because it helps us see that the same argument can be expressed in different ways. For example, the following three sentences formulate the argument that was given in standard form above.

Socrates is mortal, since all men are mortal, and Socrates is a man.

All men are mortal, so Socrates is mortal, because he is a man.

All men are mortal, and Socrates is a man, which goes to show that Socrates is mortal.

More important, by putting arguments into standard form, we perform the most obvious, and in some ways most important, step in the analysis of an argument: the identification of its premises and conclusion.

### EXERCISE III

Identify which of the following sentences expresses an argument. For each that does, (1) circle the argument marker (or markers), (2) indicate whether it is a reason marker or a conclusion marker, and (3) restate the argument in standard form.

1. Since Chicago is north of Boston, and Boston is north of Charleston, Chicago is north of Charleston.
2. Toward evening, clouds formed and the sky grew darker; then the storm broke.
3. Texas has a greater area than Topeka, and Topeka has a greater area than the Bronx Zoo, so Texas has a greater area than the Bronx Zoo.
4. Both houses of Congress may pass a bill, but the president may still veto it.
5. Other airlines will carry more passengers, because United Airlines is on strike.
6. Since Jesse James left town, taking his gang with him, things have been a lot quieter.
7. Things are a lot quieter, because Jesse James left town, taking his gang with him.
8. Witches float because witches are made of wood, and wood floats.
9. The hour is up, so you must hand in your exams.
10. Joe quit, because his boss was giving him so much grief.

## A PROBLEM AND SOME SOLUTIONS

After identifying an argument and putting it in standard form, the natural question to ask is this: Is the argument any good? If the argument is used for justification, then we can reformulate the question like this: Do the argument's premises provide a good reason to believe its conclusion?

This question will occupy us in several later chapters, but some simple examples should already be clear. Imagine that you want to buy a house, and your real estate agent shows you a particular one that looks pretty good. Then the agent tells you that this house will double in value over the next ten years. You ask, "How do you know?" The agent argues, "All of the house values in this neighborhood will double over the next ten years, so this one will double, too." Notice that this argument *does* give you a reason to believe that the conclusion is true *if* you have a reason to believe that its premise is true. However, if you have no reason to believe the premise, then the argument gives you no reason at all to believe its conclusion. In short, the argument is no good without a reason to believe its premise.

How can the real estate agent solve this problem? He needs to provide an argument for the premise, so next he argues, "All of the house values in this *city* will double over the next ten years, so all of the house values in this *neighborhood* will double over the next ten years." And if you question this new premise, he can go on to give an argument for it: "All of the house values in this *state* will double over the next ten years, so all of the house values in this *city* will double over the next ten years." And so on.

Now the problem should be obvious: An argument that aims at justification is no good unless its premises are justified. However, to justify a premise, the arguer needs to give a second argument with that premise as its conclusion. But then that second argument depends on its own premises. The second argument is no good at justifying the premise in the first argument unless the second argument's premises are justified themselves. But to justify these new premises requires a third argument, and that argument will depend on its premises being justified, which will require yet another argument, and so on. The whole process of justification seems to go on forever, requiring argument after argument without end. It now looks as if every argument, to be successful, will have to be infinitely long.

This potential regress causes deep problems in theoretical philosophy, leading some philosophers to adopt total skepticism. In everyday life, however, we try to avoid these problems by relying on shared beliefs—beliefs that will not be challenged. Beyond this, we expect people to believe us when we cite information that only we possess. But there are limits to this expectation, for we all know that people sometimes believe things that are false and sometimes lie about what they know to be true. This presents a practical problem: How can we present our reasons in a way that does not produce

just another demand for an argument—a demand for more reasons? Here we tend to use three main strategies:

1. *Assuring*: Indicating that there are backup reasons even though we are not giving them fully right now.
2. *Guarding*: Weakening our claims so that they are less subject to attack.
3. *Discounting*: Anticipating criticisms and dismissing them.

In these three ways we build a defensive perimeter around our premises. Each of these defenses is useful, but each can also be abused.

## ASSURING

When will we want to give assurances about some statement we have made? If we state something that we know everyone believes, assurances are not necessary. For that matter, if everyone believes something, we may not even state it at all; we let others fill in this step in the argument. We offer assurances when we think that someone might doubt or challenge what we say.

There are many ways to give assurances. Sometimes we cite authorities:

Doctors agree . . .

Recent studies have shown . . .

An unimpeachable source close to the White House says . . .

It has been established that . . .

Here we indicate that authorities have these reasons without specifying what their reasons are. We merely indicate that good reasons exist, even if we ourselves cannot—or choose not to—spell them out. When the authority cited can be trusted, this is often sufficient, but authorities often can and should be questioned. This topic will be discussed at greater length in Chapter 15.

Another way to give assurances is to comment on the strength of our own belief:

I'm certain that . . .

I'm sure that . . .

I can assure you that . . .

I'm not kidding. . . .

Over the years, I have become more and more convinced that . . .

Again, when we use these expressions, we do not explicitly present reasons, but we conversationally imply that there are reasons that back our assertions.

A third kind of assurance abuses the audience:

Everyone with any sense agrees that . . .

Of course, no one will deny that . . .  
It is just common sense that . . .  
There is no question that . . .  
Nobody but a fool would deny that . . .

These assurances not only do not give any reason; they also suggest that there is something wrong with you if you ask for a reason. We call this the *trick of abusive assurances*.

Just as we can give assurances that something is true, we can also give assurances that something is false. For example,

It is no longer held that . . .  
It is wholly implausible to suppose that . . .  
No intelligent person seriously maintains that . . .  
You would have to be pretty dumb to think that . . .

The last three examples clearly involve abusive assurances.

Although many assurances are legitimate, we as critics should always view assurances with some suspicion. People tend to give assurances only when they have good reasons to do so. Yet assuring remarks often mark the weakest parts of the argument, not the strongest. If someone says “I hardly need argue that . . .,” it is often useful to ask why she has gone to the trouble of saying this. When we distrust an argument—as we sometimes do—this is precisely the place to look for weakness. If assurances are used, they are used for some reason. Sometimes the reason is a good one. Sometimes, however, it is a bad one. In honest argumentation, assurances save time and simplify discussion. In a dishonest argument, they are used to paper over cracks.

## GUARDING

Guarding represents a different strategy for protecting premises from attack. We reduce our claim to something less strong. Thus, instead of saying “all,” we say “many.” Instead of saying something straight out, we use a qualifying phrase, such as “it is likely that . . .” or “it is very possible that. . . .” Law school professors like the phrase “it is arguable that. . . .” This is wonderfully noncommittal, for it does not indicate how strong the argument is, yet it does get the statement into the discussion.

Broadly speaking, there are three main ways of guarding what we say:

1. Weakening the *extent* of what has been said: retreating from “all” to “most” to “a few” to “some,” and so on.
2. Introducing *probability* phrases such as “It is virtually certain that . . .,” “It is likely that . . .,” “It might happen that . . .,” and so on.
3. Reducing our *level of commitment*: moving from “I know that . . .” to “I

believe that . . ." to "I suspect that . . .," and so on.

Such terms guard premises when they are used in place of stronger alternatives. "Madison probably quit the volleyball team" is weaker than "She definitely quit" but stronger than "She could have quit." Thus, if the context makes one expect a strong claim, such as "I know she quit," then it is guarding to say, "She probably quit." In contrast, if the context is one of speculating about who might have quit the team, then it is not guarding to say, "She probably quit." That is a relatively strong claim when others are just guessing. Thus, you need to pay careful attention to the context in order to determine whether a term has the function of guarding. When a term is used for guarding, you should be able to specify a stronger claim that the guarding term replaces and why that stronger term would be expected in the context.

Guarding terms and phrases are often legitimate and useful. If you want to argue that a friend needs fire insurance for her house, you do not need to claim that her house *will* burn down. All you need to claim is that there is a significant *chance* that her house will burn down. Your argument is better if you start with this weaker premise, because it is easier to defend and it is enough to support your conclusion.

If we weaken a claim sufficiently, we can make it completely immune to criticism. What can be said against a remark of the following kind: "There is some small chance that perhaps a few politicians are honest on at least some occasions"? You would have to have a *very* low opinion of politicians to deny this statement. On the other hand, if we weaken a premise too much, we pay a price. The premise no longer gives strong support to the conclusion.

The goal in using guarding terms is to find a middle way: We should weaken our premises sufficiently to avoid criticism, but not weaken them so much that they no longer provide strong enough evidence for the conclusion. Balancing these factors is one of the most important strategies in making and criticizing arguments.

Just as it was useful to zero in on assuring terms, so it is also useful to keep track of guarding terms. One reason is that, like assuring terms, guarding terms are easily corrupted. A common trick is to use guarding terms to *insinuate* things that cannot be stated explicitly in a conversation. Consider the effect of the following remark: "Perhaps the secretary of state has not been candid with the Congress." This does not actually say that the secretary of state has been less than candid with the Congress, but, by the rule of Relevance, clearly suggests it. Furthermore, it suggests it in a way that is hard to combat.

A more subtle device for corrupting guarding terms is to introduce a statement in a guarded form and then go on to speak as if it were not guarded at all.

Perhaps the secretary of state has not been candid with the Congress. Of course, he has a right to his own views, but this is a democracy where officials are accountable to Congress. It is time for him to level with us.

The force of the guarding term “perhaps” that begins this passage disappears at the end, where it is taken for granted that the secretary of state has not been candid. This can be called *the trick of the disappearing guard*.

What is commonly called *hedging* is a sly device that operates in the opposite direction from our last example. With hedging, one shifts ground from a strong commitment to something weaker. Things, as they say, get “watered down” or “taken back.” Strong statements made at one stage of an argument are later weakened without any acknowledgment that the position has thereby been changed in a significant way. A promise to *pass* a piece of legislation is later whittled down to a promise to *bring it to a vote*.

## DISCOUNTING

The general pattern of discounting is to cite a possible criticism in order to reject it or counter it. Notice how different the following statements sound:

The ring is beautiful, but expensive.

The ring is expensive, but beautiful.

Both statements assert the same facts—that the ring is beautiful and that the ring is expensive. Both statements also suggest that there is some opposition between these facts. Yet these statements operate in different ways. We might use the first as a reason for *not* buying the ring; we can use the second as a reason *for* buying it. The first sentence acknowledges that the ring is beautiful, but overrides this by pointing out that it is expensive. In reverse fashion, the second statement acknowledges that the ring is expensive, but overrides this by pointing out that it is beautiful. Such assertions of the form “*A* but *B*” thus have four components:

1. The assertion of *A*
2. The assertion of *B*
3. The suggestion of some opposition between *A* and *B*
4. The indication that the truth of *B* is more important than the truth of *A*

The word “but” thus discounts the statement that comes before it in favor of the statement that follows it.

“Although” is also a discounting connective, but it operates in reverse fashion from the word “but.” We can see this, using the same example:

Although the ring is beautiful, it is expensive.

Although the ring is expensive, it is beautiful.

Here the statement following the word “although” is discounted in favor of the connected statement.

A partial list of terms that typically function as discounting connectives includes the following conjunctions:

although	even if	but	nevertheless
though	while	however	nonetheless
even though	whereas	yet	still

These terms are not always used to discount. The word “still,” for example, is used for discounting in (a) “He is sick; still, he is happy” but not in (b) “He is still happy” (or “Sit still”). We can tell whether a term is being used for discounting by asking whether the sentence makes sense when we substitute another discounting term: It makes sense to say, “He is sick, but he is happy.” It makes no sense to say, “He is but happy.” It is also illuminating to try to specify the objection that is being discounted. If you cannot say which objection is discounted, then the term is probably not being used for discounting.

The clearest cases of discounting occur when we are dealing with facts that point in different directions. We discount the facts that go against the position we wish to take. But discounting is often more subtle than this. We sometimes use discounting to block certain conversational implications of what we have said. This comes out in examples of the following kind:

Jones is an aggressive player, but he is not dirty.

The situation is difficult, but not hopeless.

The Republicans have the upper hand in Congress, but only for the time being.

A truce has been declared, but who knows for how long?

Take the first example. There is no opposition between Jones being aggressive and his not being dirty. Both would be reasons to pick Jones for our team. However, the assertion that Jones is aggressive might *suggest* that he is dirty. The “but” clause discounts this suggestion without, of course, denying that Jones is aggressive.

The nuances of discounting terms can be subtle, and a correct analysis is not always easy. All the same, the role of discounting terms is often important. It can be effective in an argument to beat your opponents to the punch by anticipating and discounting criticisms before your opponents can raise them. The proper use of discounting can also help you avoid side issues and tangents.

Still, discounting terms, like the other argumentative terms we have examined, can be abused. People often spend time discounting *weak* objections to their views in order to avoid other objections that they know are harder to counter. Another common trick is to discount objections no one would raise. This is called *attacking straw men*. Consider the following remark: “A new building would be great, but it won’t be free.” This does not actually say that the speaker’s opponents think we can build a new building for free, but it

does conversationally imply that they think this, because otherwise it would be irrelevant to discount that objection. The speaker is thus trying to make the opponents look bad by putting words in their mouths that they would never say themselves. To counter tricks like this, we need to ask whether a discounted criticism is one that really would be raised, and whether there are stronger criticisms that should be raised.

#### EXERCISE IV

For each of the numbered words or expressions in the following sentences, indicate whether it is an argument marker, an assuring term, a guarding term, a discounting term, or none of these. For each argument marker, specify what the conclusion and the reasons are, and for each discounting term, specify what criticism is being discounted and what the response to this criticism is.

1. *Although* [1] no mechanism has been discovered, *most* [2] researchers in the field agree [3] that smoking greatly increases the chances [4] of heart disease.
2. *Since* [5] historically [6] public debt leads to inflation, *I maintain* [7] that, despite [8] recent trends, inflation will return.
3. *Take it from me* [9], there hasn't been a decent center fielder since [10] Joe DiMaggio.
4. *Whatever anyone tells you* [11], there is *little* [12] to the rumor that Queen Elizabeth II will step down for [13] her son, Prince Charles.
5. The early deaths of Janis Joplin and Jimi Hendrix *show* [14] that drugs are *really* [15] dangerous.
6. *I think* [16] he is out back somewhere.
7. *I think* [17], *therefore* [18] I am.
8. *I concede* [19] that the evidence is *hopelessly* [20] weak, *but* [21] I still think he is guilty.
9. *I deny* [22] that I had *anything* [23] to do with it.
10. The wind has shifted to the northeast, *which means* [24] that snow is *likely* [25].

#### EXERCISE V

1. Construct three new and interesting examples of statements containing assuring terms, and indicate which kind of assuring it is.
2. Do the same for guarding terms, and indicate which stronger claim is being reduced in strength.
3. Do the same for discounting terms, and indicate which statement is being discounted in favor of the other.
4. Do the same for argument markers, and indicate what is presented as a reason for what.

## EVALUATIVE LANGUAGE

Arguments are often filled with evaluations, so it is important to figure out what evaluative language means. We will begin with the clearest cases of evaluative language, which occur when we say simply that something is *good* or *bad*, that some course of action is *right* or *wrong*, or that something *should* or *should not* (or *ought to* or *ought not to*) be done.

Such evaluative terms often come into play when one is faced with a choice or decision. If you are deciding which shirt to buy, and a friend tells you, “That one’s good,” your friend would normally be taken to be suggesting that you get it. A passenger who says, “That’s the wrong turn,” is telling the driver not to turn that way. Evaluative language is, in these ways, used to perform speech acts of *prescribing action*.

Evaluative language is also often used to *express emotion*. When a fan says, “That band is great,” this usually expresses admiration for their music and perhaps a desire to hear more. After a meal, someone who announces, “That was horrible,” is often expressing aversion or even disgust at the food. To say, “That’s too bad,” is often to express disappointment or sadness.

Evaluative language is also typically used to bring about certain effects. When a mother tells her son that that he ought to keep his promises, she not only prescribes that her son not lie and expresses disapproval of lying; she also standardly intends to have an effect on his behavior—she tries to get him to keep his promises. And when war protesters call a war immoral, they are normally trying to get anyone listening to join their protest or at least share their disapproval. Thus, evaluative language is used to perform conversational acts of *changing people’s behavior and feelings*.

There is still more to the meaning of evaluative language. In most cases, we call something “good” or “right” because we believe that it meets or satisfies some relevant standard, and we call something “bad” or “wrong” because we believe that it violates some relevant standard. This is, roughly, the content of evaluative claims.

On this account, calling something good or bad by itself can be fairly empty, because to say that something satisfies or violates *some* standard does not explicitly specify *which* standard is satisfied or violated. Such remarks gain content—sometimes a very rich content—by virtue of the particular standards they invoke. This explains why the word “good” can be applied to so many different kinds of things. When we say that Hondas are good cars, we are probably applying standards that involve reliability, efficiency, comfort, and so on. We call someone a good firefighter because we think the person is skilled at the tasks of a firefighter, is motivated to do those tasks, works well with other firefighters, and so on. Our standards for calling someone an ethically good person concern honesty, generosity, fairness, and so on. The standards we have for calling something a good car, a good firefighter, and an (ethically) good person have little in common. Even so,

the word “good” functions in the same way in all three cases: It invokes standards that are relevant in a given context and indicates that something adequately satisfies these standards.

Because evaluative statements invoke standards, they stand in contrast to utterances that *merely* express personal feelings. If I say that I like a particular singer, then I am expressing a personal taste. It would normally be very odd for someone to reply, “No, you don’t like that singer.” On the other hand, if I call someone a good singer (or the best singer in years), then I am going beyond expressing my personal tastes. I am saying something that others may accept or reject. Of course, the standards for judging singers may be imprecise, and they may shift from culture to culture. Still, to call someone a good singer is to evaluate that person as a singer, which goes beyond merely expressing feelings, because it invokes standards and indicates that the person in question meets them.

The words “good” and “bad” are general evaluative terms. Other evaluative terms are more restrictive in their range of application. The word “delicious” is usually used for evaluating the taste of foods; it means “good-tasting.” A *sin* is a kind of wrong action, but, more specifically, it is an action that is wrong according to religious standards. A *bargain* has a good price. An *illegal* action is one that is legally wrong. Our language contains a great many specific terms of evaluation like these. Here are a few more examples:

beautiful	dangerous	wasteful	sneaky	cute
murder	prudent	nosy	sloppy	smart

Each of these words expresses either a positive or a negative evaluation of a quite specific kind.

Positive and negative evaluations can be subtle. Consider a word like “clever.” It presents a positive evaluation in terms of quick mental ability. In contrast, “cunning” often presents a negative evaluation of someone for misusing mental abilities. It thus makes a difference which one of these words we choose. It also makes a difference where we apply them. When something is supposed to be profound and serious, it is insulting to call it merely clever. Prayers, for example, should not be clever.

Sometimes seemingly innocuous words can shift evaluative force. The word “too” is the perfect example of this. This word introduces a negative evaluation, sometimes turning a positive quality into a negative one. Compare the following sentences:

John is smart.	John is too smart.
John is honest.	John is too honest.
John is ambitious.	John is too ambitious.
John is nice.	John is too nice.
John is friendly.	John is too friendly.

The word “too” indicates an excess, and thereby contains a criticism.

The difference between an evaluative term and a descriptive term is not always obvious. To see this, consider the terms “homicide” and “murder.” The words are closely related but do not mean the same thing. “Homicide” is a descriptive term meaning “the killing of a human being.” “Murder” is an evaluative term meaning, in part at least, “the *wrongful* killing of a human being.” It takes more to show that something is a murder than it does to show that something is a homicide.

Just as it is easy to miss evaluative terms because we fail to recognize the evaluative component built into their meanings, it is also possible to interpret neutral words as evaluative because of positive or negative associations with the words. The word “nuclear,” for example, has bad connotations for some people because of its association with bombs and wars, but the word itself is purely descriptive. To call people nuclear scientists is not to say that they are bad in any way.

The test for an evaluative term then is this: Does the word explicitly say that something is good or bad (right or wrong) in a particular way? A word is not evaluative when it merely suggests evaluation in some special contexts. It counts as evaluative only if its semantic content or meaning cannot be fully explained without using clearly evaluative words.

#### EXERCISE VI

Indicate whether the following italicized terms are positively evaluative (E+), negatively evaluative (E-), or simply descriptive (D). Remember, the evaluations need not be moral evaluations.

1. Janet is an *excellent* golfer.
2. The group was playing very *loudly*.
3. The group was playing *too* loudly.
4. William was *rude* to his parents.
5. William *shouted* at his parents.
6. They mistakenly turned *right* at the intersection.
7. *Fascists* ruled Italy for almost twenty years.
8. That's a *no-no*.
9. *Bummer*.
10. Debbie *lied*.
11. Debbie *said something false*.
12. Joe *copped out*.
13. Jake is a *bully*.
14. Mary Lou was a *gold medalist*.
15. She is *sick*.
16. He suffers from a hormonal *imbalance*.

**EXERCISE VII**

For each of the following sentences, construct two others—one that reverses the evaluative force, and one that is as neutral as possible. The symbol “0” stands for neutral, “+” for positive evaluative force, and “–” for negative evaluative force. Try to make as little change as possible in the descriptive content of the sentence.

Example:

- Professor Conrad is rude.
- + Professor Conrad is uncompromisingly honest in his criticisms.
- 0 Professor Conrad often upsets people with his criticisms.

- |                              |  |
|------------------------------|--|
| 1. – Larry is a lazy lout.   | 6. – Walter is a weenie.               |
| 2. + Brenda is brave.        | 7. + Carol is caring.                  |
| 3. – Sally is a snob.        | 8. – Bill is bossy.                    |
| 4. + Bartlett is a blast.    | 9. – Oprah is opinionated              |
| 5. – George is a goody-goody | 10. – This is a Mickey Mouse exercise. |



## THE ART OF CLOSE ANALYSIS

*This chapter will largely be dedicated to a single purpose: the close and careful analysis of a speech drawn from the Congressional Record, using argumentative devices introduced in Chapter 3. The point of this chapter is to show in detail how these methods of analysis can be applied to an actual argument of some richness and complexity.*

### AN EXTENDED EXAMPLE

---

It is now time to apply all of our previously discussed notions to a genuine argument. Our example will be a debate that occurred in the House of Representatives on the question of whether there should be an increase in the allowance given to members of the House for clerical help—the so-called clerk hire allowance. The argument against the increase presented by Representative Kyl (Republican, Iowa) will be examined in detail. We will put it under an analytic microscope.

The choice of this example may seem odd, for the question of clerk hire allowance is not one of the burning issues of our time. This, in fact, is one reason for choosing it. It will be useful to begin with an example about which feelings do not run high to learn the habit of objective analysis. Later on we shall examine arguments about which almost everyone has strong feelings and try to maintain an objective standpoint even there.

The example is good for two other reasons: (1) It contains most of the argumentative devices we have listed, and (2) relatively speaking, it is quite a strong argument. This last remark may seem ironic after we seemingly tear the argument to shreds. However, in comparison to other arguments we shall examine, it stands up well.

We begin by reading through a section of the *Congressional Record* (vol. 107, part 3, March 15, 1961, pp. 4059–60) without comment:

#### CLERK HIRE ALLOWANCE, HOUSE OF REPRESENTATIVES

---

Mr. FRIEDEL. Mr. Speaker, by direction of the Committee on House Administration, I call up the resolution (H. Res. 219) to increase the basic clerk hire allowance of each Member of the House, and for other purposes, and ask for its immediate consideration.

The Clerk read the resolution as follows:

Resolved, That effective April 1, 1961, there shall be paid out of the contingent fund of the House, until otherwise provided by law, such sums as may be necessary to increase the basic clerk hire allowance of each Member and the Resident Commissioner from Puerto Rico by an additional \$3,000 per annum, and each such Member and Resident Commissioner shall be entitled to one clerk in addition to those to which he is otherwise entitled by law.

Mr. FRIEDEL. Mr. Speaker, this resolution allows an additional \$3,000 per annum for clerk hire and an additional clerk for each Member of the House and the Resident Commissioner from Puerto Rico. Our subcommittee heard the testimony, and we were convinced of the need for this provision to be made. A few Members are paying out of their own pockets for additional clerk hire. This \$3,000 is the minimum amount we felt was necessary to help Members pay the expenses of running their offices. Of course, we know that the mail is not as heavy in some of the districts as it is in others, and, of course, if the Member does not use the money, it remains in the contingent fund.

Mr. KYL. Mr. Speaker, will the gentleman yield?

Mr. FRIEDEL. I yield to the gentleman from Iowa [Mr. Kyl] for a statement.

Mr. KYL. Mr. Speaker, I oppose this measure. I oppose it first because it is expensive. I further oppose it because it is untimely.

I do not intend to belabor this first contention. We have been presented a budget of about \$82 billion. We have had recommended to us a whole series of additional programs or extensions of programs for priming the pump, for depressed areas, for the needy, for the unemployed, for river pollution projects, and recreation projects, aid to education, and many more. All are listed as "must" activities. These extensions are not within the budget. Furthermore, if business conditions are as deplorable as the newspapers indicate, the Government's income will not be as high as anticipated. It is not enough to say we are spending so much now, a little more will not hurt. What we spend, we will either have to recover in taxes, or add to the staggering national debt.

The amount of increase does not appear large. I trust, however, there is no one among us who would suggest that the addition of a clerk would not entail allowances for another desk, another typewriter, more materials, and it is not beyond the realm of possibility that the next step would then be a request for additional office space, and ultimately new buildings. Some will say, "All the Members will not use their maximum, so the cost will not be great." And this is true. If the exceptions are sufficient in number to constitute a valid argument, then there is no broad general need for this measure. Furthermore, some Members will use these additional funds to raise salaries. Competition will force all salaries upward in all offices and then on committee staffs, and so on. We may even find ourselves in a position of paying more money for fewer clerks and in a tighter bind on per person workload.

This measure proposes to increase the allowance from \$17,500 base clerical allowance to \$20,500 base salary allowance. No member of this House can tell us what this means in gross salary. That computation is almost

impossible. Such a completely absurd system has developed through the years on salary computations for clerical hire that we have under discussion a mathematical monstrosity. We are usually told that the gross allowed is approximately \$35,000. This is inaccurate. In one office the total might be less than \$35,000 and in another, in complete compliance with the law and without any conscious padding, the amount may be in excess of \$42,000. This is possible because of a weird set of formulae which determines that three clerks at \$5,000 cost less than five clerks at \$3,000. Five times three might total the same as three times five everywhere else in the world—but not in figuring clerk hire in the House.

This is an application of an absurdity. It is a violation of bookkeeping principles, accounting principles, business principles and a violation of common sense. Listen to the formula:

First, 20 percent increase of first \$1,200; 10 percent additional from \$1,200 to \$4,600; 5 percent further additional from \$4,600 to \$7,000.

Second, after applying the increases provided in paragraph 1, add an additional 14 percent or a flat \$250, whichever is the greater, but this increase must not exceed 25 percent.

Third, after applying the increases provided in both paragraphs 1 and 2, add an additional increase of 10 percent in lieu of overtime.

Fourth, after applying the increases provided in paragraphs 1, 2, and 3, add an additional increase of \$330.

Fifth, after applying the increases provided in paragraphs 1, 2, 3, and 4, add an additional increase of 5 percent.

Sixth, after applying the increases provided in paragraphs 1, 2, 3, 4, and 5, add an additional increase of 10 percent but not more than \$800 nor less than \$300 a year.

Seventh, after applying the increases provided in paragraphs, 1, 2, 3, 4, 5, and 6, add an additional increase of 7½ percent.

Eighth, after applying the increases provided in paragraphs, 1, 2, 3, 4, 5, 6, and 7, add an additional increase of 10 percent.

Ninth, after applying the increases provided in paragraphs 1, 2, 3, 4, 5, 6, 7, and 8, add an additional increase of 7½ percent.

The Disbursing Office has a set of tables to figure House salaries for office staffs and for about 900 other employees. It contains 45 sheets with 40 entries per sheet. In the Senate, at least, they have simplified the process some by figuring their base in multiples of 60, thus eliminating 11 categories. Committee staffers, incidentally, have an \$8,880 base in comparison to the House \$7,000 base limitation.

Now, Mr. Speaker, I have planned to introduce an amendment or a substitute which would grant additional clerk hire where there is a demonstrable need based on heavier than average population or “election at large” and possible other factors. But after becoming involved in this mathematical maze, I realize the folly of proceeding one step until we have corrected this situation. We can offer all kinds of excuses for avoiding a solution. We cannot offer reasonable arguments that it should not be done or that it cannot be done.

Someone has suggested that the Members of this great body prefer to keep the present program because someone back in the home district might object to the gross figures. I know this is not so. When a Representative is busy on minimum wage, or aid to education, or civil rights, such matters of housekeeping seem too picayune to merit attention. The Member simply checks the table and hires what he can hire under the provisions and then forgets the whole business. But I know the Members also want the people back home to realize that what we do here is open and frank and accurate, and that we set an example in businesslike procedures. The more we can demonstrate responsibility the greater will be the faith in Congress.

May I summarize. It is obvious that some Members need more clerical help because of large population and large land area. I have been working for some time with the best help we can get, on a measure which would take these items into consideration. Those Members who are really in need of assistance should realize that this temporary, hastily conceived proposition we debate today will probably obviate their getting a satisfactory total solution.

First, we should await redistricting of the Nation.

Second, we should consider appropriate allowance for oversize districts considering both population and total geographic area.

Finally, I hope we can develop a sound and sensible formula for computing salaries of office clerks and other statutory employees in the same category.

---

Before going any further, it will be useful to record your general reactions to this speech. Perhaps you think that on the whole Kyl gives a well-reasoned argument on behalf of his position. Alternatively, you might think that he is making a big fuss over nothing, trying to confuse people with numbers, and just generally being obnoxious. When you are finished examining this argument in detail, you can look back and ask yourself why you formed this original impression and how, if at all, you have changed your mind.

The first step in the close analysis of an argument is to go through the text, labeling the various argumentative devices we have examined. Here some abbreviations will be useful:

argument marker	M
assuring term	A
guarding term	G
discounting term	D
argumentative performative	AP
evaluative term	E (+ or -)
rhetorical device	R

The last label is a catchall for the various rhetorical devices discussed in Chapter 1, such as overstatement, understatement, irony or sarcasm, metaphor, simile, rhetorical questions, and so on.

If you want to make your analysis extra close, it is illuminating to specify which rhetorical device is deployed whenever you mark something with "R."

It is also useful to specify whether each argument marker marks a reason or a conclusion (and what the argument is), which stronger term is replaced by each guarding term marked “G,” and which objection is discounted whenever you mark a discounting term with “D.”

This simple process of labeling brings out features of an argument that could pass by unnoticed. It also directs us to ask sharp critical questions. To see this, we can look at each part of the argument in detail.

M                   Mr. KYL. Mr. Speaker [I oppose] this measure. [I oppose] it AP  
 first [because] it is expensive. I further [oppose] it [because] it is AP  
 untimely. M

This is a model of clarity. By the use of a performative utterance in the opening sentence, Kyl makes it clear that he opposes the measure. Then by twice using the argument marker “because,” he gives his two main reasons for opposing it: It is expensive and it is untimely. We must now see if he makes good on each of these claims.

The next paragraph begins the argument for the claim that the measure is expensive:

[I do not intend to belabor this first contention] We have been A  
 presented a budget of about \$82 billion. We have had rec-  
 ommended to us a whole series of additional programs or  
 extensions of programs for priming the pump, for depressed  
 areas, for the needy, for the unemployed, for river pollution  
 projects, and recreation projects, aid to education, and many  
 more. All are listed as @must@ activities. These extensions R  
 are not within the budget. Furthermore, if business condi- M  
 tions are as deplorable as the newspapers indicate, the Gov-  
 ernment’s income will not be as high as anticipated. [It is not  
 enough to say] we are spending so much now, a little more D  
 will not hurt. What we spend, we will either have to recover  
 in taxes, or add to the staggering national debt.

a. “I do not intend to belabor this first contention.” This is an example of *assuring*. The conversational implication is that the point is so obvious that little has to be said in its support. Yet there is something strange going on here. Having said that he will *not* belabor the claim that the bill is expensive, Kyl actually goes on to say quite a bit on the subject. It is a good idea to look closely when someone says that he or she is not going to do something, for often just the opposite is happening. For example, saying “I am not suggesting that Smith is dishonest” is one way of suggesting that Smith *is* dishonest. If no such suggestion is being made, why raise the issue at all?

b. Kyl now proceeds in a rather flat way, stating that the proposed budget comes to \$82 billion and that it contains many new programs and extensions of former programs. Because these are matters of public record and nobody is likely to deny them, there is no need for guarding or assuring. Kyl also claims, without qualification, that these extensions are not within the budget. This recital of facts does, however, carry an important conversational implication: Since the budget is already out of balance, any further extensions should be viewed with suspicion.

c. Putting the word “must” in quotation marks, or saying it in a sarcastic tone of voice, is a common device for denying something. The plain suggestion is that some of these measures are *not* must activities at all. Kyl here suggests that some of the items already in the budget are not necessary. He does this, of course, without defending this suggestion.

d. “Furthermore, if business conditions are as deplorable as the newspapers indicate, the Government’s income will not be as high as anticipated.” The word “furthermore” suggests that an argument is about to come. However, the following sentence as a whole is an *indicative conditional* (with the word “then” dropped out). As such, the sentence does not produce an argument, but instead provides only a pattern for an argument.

To get an argument from this pattern, one would have to assert the antecedent of the conditional. The argument would then come to this:

- (1) If business conditions are as deplorable as the newspapers indicate,  
then the Government’s income will not be as high as anticipated.
  - (2) Business conditions are as deplorable as the newspapers indicate.
- ∴(3) The Government’s income will not be as high as anticipated.

The first premise seems perfectly reasonable, so, if Kyl could establish the second premise, then he would have moved the argument along in an important way. Yet he never explicitly states that business conditions are so deplorable. All he says is that “the newspapers *indicate*” this. Moreover, this appeal to authority (see Chapter 15) does not mention any specific newspaper, so he does not endorse any specific authority. Still, Kyl never questions what the newspapers claim, and it would be misleading to bring up these newspaper reports without questioning them if he thought they were way off the mark. So Kyl does seem to have in mind something like the arguments (1)–(3).

e. “It is not enough to say we are spending so much now, a little more will not hurt.” The opening phrase is, of course, used to deny what follows it. Kyl is plainly rejecting the argument that, since we are spending so much now, a little more will not hurt. Yet his argument has a peculiar twist, for who would come right out and make such an argument? If you stop to think for a minute, it should be clear that nobody would want to put it that way. An opponent, for example, would use quite different phrasing. He might say something like this: “Considering the large benefits that will flow from this measure, it is more than worth the small costs.” What Kyl has done is attribute a bad argument to his opponents and then reject it in an indignant

tone. This is a common device, and when it is used, it is often useful to ask whether anyone would actually argue or speak in the way suggested. When the answer to this question is no, as it often is, we have what was called “the trick of discounting straw men” in Chapter 3 (see also Chapter 17). In such cases, it is useful to ask what the speaker’s opponent would have said instead. This leads to a further question: Has the arguer even addressed himself to the *real* arguments of his opponents?

So far, Kyl has not addressed himself to the first main point of his argument: that the measure is *expensive*. This is not a criticism, because he is really making the preliminary point that the matter of expense is significant. Here he has stated some incontestable facts—for example, that the budget is already out of balance. Beyond this he has indicated, with varying degrees of strength, that the financial situation is grave. It is against this background that the detailed argument concerning the cost of the measure is actually presented in the next paragraph.

The amount of increase does not appear large. (I trust,  
D      however) (there is no one among us who would suggest) A  
that the addition of a clerk would not entail allowances for  
another desk, another typewriter, more materials, (and it is  
not beyond the realm of possibility) that the next step would  
then be a request for additional office space, and ultimately  
D      new buildings. (Some will say), “All the Members will not  
use their maximum, so the cost will not be great.” And this  
is true. If the exceptions are sufficient in number to consti-  
tute a valid argument, then there is no broad general need  
M      for this measure. (Furthermore) (some) Members will use these G  
additional funds to raise salaries. Competition will force all  
salaries upward in all offices and then on committee staffs,  
G      and so on. We (may) even find ourselves in a position of  
paying more money for fewer clerks and in a tighter bind on  
per person workload.

a. “The amount of increase does not appear large.” Words like “appear” and “seem” are sometimes used for guarding, but we must be careful not to apply labels in an unthinking way. The above sentence is the beginning of a discounting argument. As soon as you hear this sentence, you can feel that a word like “but” or “however” is about to appear. Sure enough, it does.

b. “I trust, however, there is no one among us who would suggest that the addition of a clerk would not entail allowances for another desk, another typewriter, more materials. . . .” This is the beginning of Kyl’s argument that is intended to rebut the argument that the increase in expenses will not be large. Appearances to the contrary, he is saying, the increase *will* be large. He then ticks off some additional expenses that are entailed by hiring new

clerks. Notice that the whole sentence is covered by the assuring phrase “I trust . . . there is no one among us who would suggest. . . .” This implies that anyone who would make such a suggestion is merely stupid. But the trouble with Kyl’s argument so far is this: He has pointed out genuine additional expenses, but they are not, after all, very large. It is important for him to get some genuinely large sums of money into his argument. This is the point of his next remark.

c. “. . . and it is not beyond the realm of possibility that the next step would then be a request for additional office space, and ultimately new buildings.” Here, at last, we have some genuinely large sums of money in the picture, but the difficulty is that the entire claim is totally guarded by the phrase “it is not beyond the realm of possibility.” There are very few things that *are* beyond the realm of possibility. Kyl’s problem, then, is this: There are certain additional expenses that he can point to without qualification, but these tend to be small. On the other hand, when he points out genuinely large expenses, he can only do so in a guarded way. So we are still waiting for a proof that the expense will be large. (Parenthetically, it should be pointed out that Kyl’s prediction of new buildings actually came true.)

d. “Some will say, ‘All the Members will not use their maximum, so the cost will not be great.’ And this is true. If the exceptions are sufficient in number to constitute a valid argument, then there is no broad general need for this measure.” This looks like a “trick” argument, and for this reason alone it demands close attention. The phrase “some will say” is a standard way of beginning a discounting argument. This *is*, in fact, a discounting argument, but its form is rather subtle. Kyl cites what some will say, and then he adds, somewhat surprisingly: “And this is true.” To understand what is going on here, we must have a good feel for conversational implication. Kyl imagines someone reasoning in the following way:

All the Members will not use their maximum.

So, the cost will not be great.

Therefore, we should adopt the measure.

Given the same starting point, Kyl tries to derive just the *opposite* conclusion along the following lines:

All the Members will not use their maximum.

If the exceptions are not sufficient, then the cost will be too great.

But if the exceptions are sufficient, there is no broad general need for this measure.

Therefore, whether it is expensive or not, we should reject this measure.

In order to get clear about this argument, we can put it into schematic form:

*Kyl’s argument:*

If (1) the measure is expensive, then *reject* it.

If (2) the measure is inexpensive, then, because that shows there is no general need, *reject* it.

*The opposite argument:*

If (1) the measure is inexpensive, then *accept* it.

If (2) the measure is expensive, then, because that demonstrates a general need, *accept* it.

When the arguments are spread out in this fashion, it should be clear that they have equal strength. Both are no good. The question that must be settled is this: Does a genuine need exist that can be met in an economically sound manner? If there is no need for the measure, then it should be rejected, however inexpensive. Again, if there is a need, then some expense is worth paying. The real problem is to balance need against expense and then decide on this basis whether the measure as a whole is worth adopting.

Kyl's argument is a sophistry, because it has no tendency to answer the real question at hand. A *sophistry* is a clever but fallacious argument intended to establish a point through trickery. Incidentally, it is one of the marks of a sophistical argument that, though it may baffle, it almost never convinces. We think that few readers will have found this argument persuasive even if they cannot say exactly what is wrong with it. The appearance of a sophistical argument (or even a complex and tangled argument) is a sign that the argument is weak. Remember, when a case is strong, people usually argue in a straightforward way.

e. "Furthermore, some Members will use these additional funds to raise salaries. Competition will force all salaries upward in all offices and then on committee staffs, and so on." The word "furthermore" signals that further reasons are forthcoming. Here Kyl returns to the argument that the measure is more expensive than it might appear at first sight. Although Kyl's first sentence is guarded by the term "some," he quickly drops his guard and speaks in an unqualified way about *all* salaries in *all* offices. Yet the critic is bound to ask whether Kyl has any right to make these projections. Beyond this, Kyl here projects a *parade of horrors*. (See Chapter 13.) He pictures this measure leading by gradual steps to quite disastrous consequences. Here the little phrase "and so on" carries a great burden in the argument. Once more, we must simply ask ourselves whether these projections seem reasonable.

f. "We may even find ourselves in a position of paying more money for fewer clerks and in a tighter bind on per person workload." Once more, the use of a strong guarding expression takes back most of the force of the argument. Notice that if Kyl could have said straight out that the measure *will* put us in a position of paying more money for fewer clerks and in a tighter bind on per-person workload, that would have counted as a very strong objection. You can hardly do better in criticizing a position than showing that it will have just the opposite result from what is intended. In fact, however, Kyl has not established this; he has only said that this is something that we "may even find."

Before we turn to the second half of Kyl's argument, which we shall see in a moment is much stronger, we should point out that our analysis has not been entirely fair. Speaking before the House of Representatives, Kyl is in an *adversarial* situation. He is not trying to prove things for all time; rather, he is responding to a position held by others. Part of what he is doing is *raising objections*, and a sensitive evaluation of the argument demands a detailed understanding of the nuances of the debate. But even granting this, it should be remembered that objections themselves must be made for good reasons. The problem so far in Kyl's argument is that the major reasons behind his objections have constantly been guarded in a very strong way.

Turning now to the second part of Kyl's argument—that the measure is untimely—we see that he moves along in a clear and direct way with little guarding.

This measure proposes to increase the allowance from \$17,500 base clerical allowance to \$20,500 base salary allowance. No member of this House can tell us what this means in gross salary. That computation is almost impossible. Such a completely absurd system has developed through the years on salary computations for clerical hire that we have under discussion a mathematical monstrosity. We are usually told that the gross allowed is approximately \$35,000. This is inaccurate. In one office the total might be less than \$35,000 and in another, in complete compliance with the law and without any conscious padding, the amount may be in excess of \$42,000. This is possible because of a weird set of formulae which determines that three clerks at \$5,000 cost less than five clerks at \$3,000. Five times three might total the same as three times five everywhere else in the world—but not in figuring clerk hire in the House. This is an application of an absurdity. It is a violation of bookkeeping principles, accounting principles, business principles and a violation of common sense. Listen to the formula.

The main point of the argument is clear enough: Kyl is saying that the present system of clerk salary allowance is utterly confusing, and this matter should be straightened out before *any* other measures in this area are adopted. There is a great deal of negative evaluation in this passage. Notice the words and phrases that Kyl uses:

- a completely absurd system
- weird set of formulae
- violation of common sense

mathematical monstrosity

an absurdity

There is also a dash of irony in the remark that five times three might total the same as three times five everywhere else in the world, but not in figuring clerk hire in the House. Remember, there is nothing wrong with using negative evaluative and expressive terms if they are deserved. Looking at the nine-step formula in Kyl's speech, you can decide for yourself whether he is on strong grounds in using this negative language.

Now, Mr. Speaker, I have planned to introduce an amendment or a substitute which would grant additional clerk hire where there is a demonstrable need based on heavier than average population or "election at large" and possible other factors.

a. This passage discounts any suggestion that Kyl is unaware that a genuine problem does exist in some districts. It also indicates that he is willing to do something about it.

b. The phrase "and possible other factors" is not very important, but it seems to be included to anticipate other reasons for clerk hire that should at least be considered.

D ————— (But) after becoming involved in this mathematical maze I E—  
A ————— realize the (folly) of proceeding one step until we have corrected this situation. E—

a. Here Kyl clearly states his reason for saying that the measure is untimely. Notice that the reason offered has been well documented and is not hedged in by qualifications.

b. The phrases "mathematical maze" and "folly" are again negatively evaluative.

We can offer all kinds of (excuses) for avoiding a solution. E—  
(We cannot offer reasonable arguments that) it should not be A  
done or that it cannot be done.

Notice that the first sentence ridicules the opponents' arguments by calling them *excuses*, a term with negative connotations. The second sentence gives assurances that such a solution can be found.

Someone has suggested that the Members of this great body prefer to keep the present program because someone back in the home district might object to the gross figures. (I know) A  
this is not so. When a Representative is busy on minimum

wage, or aid to education, or civil rights, such matters of housekeeping seem too picayune to merit attention. The Member simply checks the table and hires what he can hire under the provisions and then forgets the whole business.

D (But I know) the Members also want the people back home to realize that what we do here is open and frank and accurate and that we set an example in businesslike procedures. The more we can demonstrate responsibility the greater will be the faith in Congress. A E+

a. Once more the seas of rhetoric run high. Someone (though not Kyl himself) has suggested that the members of the House wish to conceal information. He disavows the very thought that he would make such a suggestion by the sentence "I know this is not so." All the same, he has gotten this suggestion into the argument.

b. Kyl then suggests another reason why the members of the House will not be concerned with this measure: It is "too picayune." The last two sentences rebut the suggestion that it is too small to merit close attention. Even on small matters, the more the House is "open and frank and accurate," the more it will "set an example in businesslike procedures" and thus "demonstrate responsibility" that will increase "the faith in Congress." This is actually an important part of Kyl's argument, for presumably his main problem is to get the other members of the House to take the matter seriously.

M May I summarize. It is obvious that some Members need more clerical help because of large population and large land area. I have been working for some time with the best help we can get on a measure which would take these items into consideration. Those Members who are really in need A E-

A of assistance should realize that this temporary hastily conceived proposition we debate today will probably obviate G their getting a satisfactory total solution.

a. This is a concise summary. Kyl once more assures the House that he is aware that a genuine problem exists. He also indicates that he is working on it.

b. The phrase "temporary, hastily conceived proposition we debate today" refers back to his arguments concerning untimeliness.

c. The claim that "it will probably obviate their getting a satisfactory total solution" refers back to the economic argument. Notice, however, that, as before, the economic claim is guarded by the word "probably."

E+ First, we should await redistricting of the Nation.

E+ Second, we should consider appropriate allowance for oversize districts considering both population and total geographic area.

Finally, I hope we can develop a sound and sensible formula for computing salaries of office clerks and other statutory employees in the same category.

E+

This is straightforward except that a new factor is introduced: We should await redistricting of the nation. This was not mentioned earlier in the argument, and so seems a bit out of place in a summary. Perhaps the point is so obvious that it did not need any argument to support it. On the other hand, it is often useful to keep track of things that are smuggled into the argument at the very end. If redistricting was about to occur in the *near* future, this would give a strong reason for delaying action on the measure. Because the point is potentially so strong, we might wonder why Kyl has made so little of it. Here, perhaps, we are getting too subtle.

Now that we have looked at Representative Kyl's argument in close detail, we can step back and notice some important features of the argument as a whole. In particular, it is usually illuminating to notice an argument's *purpose, audience, and standpoint*.

First, Kyl's overall *purpose* is clear. As his opening sentence indicates, he is presenting an argument intended to justify his opposition to an increase in the clerk hire allowance. Virtually everything he says is directed toward this single goal. In other cases, arguers pursue multiple goals, and sorting things out can be a complex matter. Sometimes it is hard to tell what an argument is even intended to establish. This is usually a sign that the person presenting the argument is confused or, perhaps, trying to confuse his audience.

Second, Kyl's argument is addressed to a specific *audience*. He is not speaking to an enemy of the United States who would love to see our government waste its money. Nor is he speaking to clerks or to those U.S. citizens who might be hired as clerks if the clerk hire allowance were raised. He is presenting his argument to other representatives in Congress. He is trying to show this group that they and he have reasons to oppose this increase in the clerk hire allowance. His task, then, is to present reasons that *they* accept—or should accept—for rejecting an increase in the clerk hire allowance.

Third, Kyl not only addresses his argument to a particular audience, he also adopts a particular *standpoint* to it. Good arguments are usually presented not only *to* specific audiences but also *from* particular standpoints. Kyl's standpoint is clear and powerful. He puts himself across as a tough-minded, thoroughly honest person who is willing to stand up against majority opinion. This, in fact, may be an accurate representation of his character, but by adopting this standpoint he gains an important argumentative advantage: He suggests that those who disagree with him are a bit soft-minded, not altogether candid, and, anyway, mere tools of the Democratic majority that runs the Congress. By adopting this stance, Kyl casts his opponents in a light that is hardly flattering.

By specifying the purpose, audience, and standpoint of an argument, we get a clearer sense of what the argument needs to accomplish in order to succeed in its goals. By looking closely at special words in the argument, as well as at what is conversationally implied, we get a better idea of how the argument is supposed to achieve its goals. All of this together helps us understand the

argument. It will sometimes remain unclear how well the argument succeeds. It will always require care and skill to apply these methods. Still, the more you practice, the more you will be able to understand arguments.

### EXERCISE I

Read the following passage. Then, for each of the numbered expressions, either answer the corresponding question or label the main argumentative move, if any, using these abbreviations:

- M = argument marker
- A = assuring term
- G = guarding term
- D = discounting term
- E- = negative evaluative term
- E+ = positive evaluative term
- R = rhetorical device
- N = none of the above

This letter to the editor appeared in *The Dartmouth* on September 23, 1992, although references to the author's college have been removed. The author was president of the student assembly and a member of a single-sex fraternity at the time.

### GREEKS SHOULD BE CO-ED

by Andrew Beebe

For some time now, people have been asking the question "Why should the Greek [fraternity and sorority] system go co-ed?" To them, *I pose an answer* [1] in a question, "*Why not?*" [2]

*Learning in college extends beyond the classrooms, onto the athletic fields, into the art studios, and into our social environs.* [3] *In fact* [4], *some* [5] say that *most* [6] of what we learn at college comes from interaction with people and ideas during time spent outside of the lecture halls. The concept of segregating students in their social and residential environments by gender directly contradicts the *ideals* [7] of a college experience. This is *exactly* [8] what the fraternity and sorority system does.

With all the *benefits* [9] of a small, closely-bonded group, the potential for strong social education would seem *obvious* [10]. *But* [11] is it *fair* [12] for us to remove the other half of our community from that education? [13] In many colleges, this voluntary segregation exists in fraternities and sororities.

From the planning of a party or involvement in student activities to the sharing of living and recreational space, the fraternity and sorority system is a social environment *ripe* [14] with *educational potential* [15]. The idea that women and men would receive as complete an experience from these environments while *virtually* [16] separated is *implausible* [17].

Source: Dartmouth: Letter to the Editor. "Greeks Should be Co-Ed" by Andrew Beebe, September 23, 1992. Used by permission of Andrew Beebe.

*But [18] what do women and men learn from one another that they don't already know? [19] Problems in gender relations between all ages prove [20] that our society is plagued by gender-based prejudice [21]. Since [22] prejudice is the ignorance of one group by another [23], it will best be addressed by education. The question then [24] becomes: Which way is best to educate one another?*

Sexism, homophobia, date rape, eating disorders, and other social *problems* [25] are often [26] connected to gender-relation issues. As campus experience shows [27], we have a long way to go in combating these problems. Defenders of fraternities and sororities may [28] argue that they do not, solely by nature of being single sex, promote sexism or other prejudices. But [29], if we can recognize that these problems exist in our society, it is not important to find the blame, but [30] rather to offer a solution. It is clear [31] that separating people by gender is not the right [32] way to promote better [33] understanding between the sexes. To the contrary, bringing different people together is the only way prejudice, no matter what the cause (or result) may be [34], can be overcome.

Acknowledging that breaking down walls of separation may [35] help foster better understanding, it is important to look at what might [36] change for the worse. There would be some [37] obvious [38] logistical changes in rush, pledging, relationships with national organization, and house leadership. But [39] where are the real consequences? [40] Men could [41] still cultivate strong bonds with other men. Women could [42] still bond with other women. The difference is that there would be a well-defined [43] environment where men and women could [44] create strong, lasting bonds and friendships between one another.

There are many more benefits [45] to a co-ed system than there are sacrifices [46]. Men and women could share the responsibilities of running what is now a predominantly [47] male-controlled social structure. First-year men and women could interact with older students in a social environment beyond the classroom or the dining halls. People in a co-ed system could find a strong support group that extends beyond their own sex. With these advantages [48] and more, it is clear [49] that the all-co-ed system offers everything found in a single-sex organization and more. Although [50] there are some [51] minor sacrifices to be made, they are insignificant in comparison to the gain [52] for all.

College is the last place we want to isolate ourselves. The entire idea of the "holistic education" is based on [53] expanding our knowledge, not separating ourselves from one another. Our fraternity and sorority system includes many [54] different types of students. So [55] why should some houses refuse women simply because [56] they are women? Why do some houses refuse men solely because [57] they are men? The only solution is desegregation of the fraternity and sorority system. After all [58], when it comes to challenging one another to learn, what are we afraid of? [59]

#### QUESTIONS:

- [1]: Is this sentence an explicit performative?
- [2]: Explain the difference between asking "Why?" and asking "Why not?" in this context.
- [3]: Why does the author begin with this point?

(continued)

- [4]–[12]: Write labels.
- [13]: What is the expected answer to this rhetorical question?
- [14]: What kind of rhetorical device is this? What is its point?
- [15]–[18]: Write labels.
- [19]: Who is supposed to be asking this question?
- [20]–[22]: Write labels.
- [23]: What is the point of this definition?
- [24]–[33]: Write labels.
- [34]: Why does this author add this dependent clause?
- [35]–[39]: Write labels.
- [40]: What does this question imply in this context?
- [41]–[58]: Write labels.
- [59]: What is the expected answer to this rhetorical question?

### EXERCISE II

Read the following passage from *The Washington Post* (November 25, 1997), page A19. Then, for each of the numbered expressions, label the main argumentative move, if any, using the same abbreviations as in Exercise I:

#### **A PIECE OF “GOD’S HANDIWORK”**

*by Robert Redford*

Just over a year ago, President Clinton created the Grand Staircase-Escalante National Monument to [1] protect [2] once and for all some [3] of Utah’s extraordinary red rock canyon country. In response to [4] plans of the Dutch company Andalex to mine coal on the Kaiparowits Plateau, President Clinton used his authority under the Antiquities Act to establish the new monument, setting aside for protection what he described as “some of the most remarkable land in the world.” *I couldn’t agree more.* [5] For over two decades, many have fought battle after battle [6] to keep mining conglomerates from despoiling [7] the unique treasures [8] of this stunning red rock canyon country. Now [9], we thought at least some of it was safe.

Not so. Shocking [10] as it sounds, Clinton’s Bureau of Land Management (BLM) has approved oil drilling within the monument. BLM has given Conoco Inc., a subsidiary of the corporate giant DuPont, permission to drill for oil and gas in the heart [11] of the new monument. You may [12] wonder, as I do, *how can this happen?* [13] Wasn’t the whole purpose of creating the monument to preserve its colorful cliffs, sweeping arches and other extraordinary resources [14] from large-scale mineral development? Didn’t the president say he

Source: Robert Redford, “A Piece of ‘God’s Handiwork,’” *Washington Post*, November 25, 1997, p. A19. Used by permission.

was saving [15] these lands from mining companies for our children and grandchildren?

The BLM says *its hands are tied*. [16] Why? Because [17] these lands were set aside subject to “valid existing rights,” and Conoco has a lease that gives it the right to drill. Sure [18] Conoco has a lease—more than one, *in fact* [19]—but [20] those leases were originally issued without sufficient environmental study or public input. As a result [21], none of them conveyed a valid right to drill. What’s more [22], in deciding to issue a permit to drill now, the BLM did not conduct a full analysis of the environmental impacts of drilling in these incomparable lands, but instead [23] determined there would be no significant environmental harm on the basis of an abbreviated review that didn’t even look at drilling on the other federal leases.

Sounds like [24] Washington double-speak [25] to me. I’ve spent considerable time on these extraordinary lands for years, and I know [26] that an oil rig in their midst would have a major impact. What’s more [27], Conoco wants to drill a well to find oil. Inevitably [28], more rigs, more roads, new pipelines, toxic [29] wastes and bright lights would follow to get the oil out. The BLM couldn’t see this, but [30] the U.S. Fish and Wildlife Service and the Environmental Protection Agency did. Both of those agencies recognized [31] the devastating [32] effects extensive oil drilling would have on this area and urged the BLM to refuse to allow it, *in order to* [33] protect the monument.

Maybe [34] the problem [35] comes from giving management responsibility for this monument to the BLM. This is the BLM’s first national monument; almost [36] all the others are managed by the National Park Service. The Park Service’s mission is to protect the resources [37] under its care while the bureau has always sought to accommodate economic uses of those under its. Even so [38], the BLM seemed [39] to be getting off to a good [40] start by enlisting broad [41] public involvement in developing a management plan for the area. Yet [42] the agency’s decision to allow oil drilling in the monument completely undercuts [43] this process just as it is beginning.

What we’re talking about is, in the words of President Clinton, a small piece of “God’s handiwork.” Almost [44]  $4\frac{1}{2}$  million acres of irreplaceable red rock wilderness remain outside the monument. Let us at least protect what is within it. The many roadless [45] areas within the monument should [46] remain so—protected as wilderness. The monument’s designation means little if [47] a pattern of exploitation is allowed to continue.

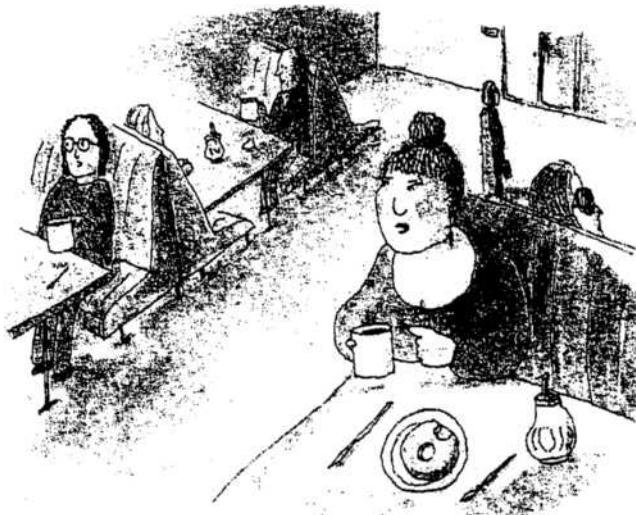
Environmentalists—including the Southern Utah Wilderness Alliance, the Natural Resources Defense Council, and the Wilderness Society—appealed BLM’s decision to the Interior Department’s Board of Land Appeals. This appeal, however [48], was rejected earlier this month. This is a terrible mistake [49]. We shouldn’t be drilling in our national monuments. Period. As President Clinton said when dedicating the new monument, “Sometimes progress is measured in mastering frontiers, but sometimes [50] we must measure progress in protecting frontiers for our children and children to come.”

Allowing drilling to go forward in the Grand Staircase-Escalante Monument would permanently stain what might otherwise have been a defining legacy of the Clinton presidency.

**EXERCISE III**

Read the following advertisement from *Equal Exchange* (Copyright © 1997, 1998, 1999). For each of the numbered expressions, label the main argumentative move, if any, using the same abbreviations as in Exercise I. Then state what you take to be the central conclusions and premises. What criticisms, if any, do you have of this argument?

It *may* [1] be a little early in the morning to bring this up, *but* [2] *if* [3] you buy coffee from *large corporations* [4], you are inadvertently maintaining the system which keeps small farmers *poor* [5] *while* [6] *lining the pockets* [7] of rich corporations. *By* [8] choosing Equal Exchange coffee, you *can* [9] help to make a change. We believe in trading directly with small farming cooperatives at mutually agreed-upon prices with a fixed minimum rate. *Then* [10], *should* [11] the coffee market decline, the farmers are still guaranteed a *fair* [12] price. *So* [13] have a cup of Equal Exchange Coffee and make a small farmer *happy* [14]. *Of course* [15], your decision to buy Equal Exchange need not be completely altruistic. *For* [16] we take as much pride in refining the taste of our *gourmet* [17] coffees as [18] we do in *helping* [19] the farmers who produce them. *For* [20] more information about Equal Exchange or to order our line of gourmet, organic, and shade-grown coffee directly, call 1 800 406 8289.



*"Excuse me waiter, there's the blood and misery  
of a thousand small farmers in my coffee."*

From *Equal Exchange*. Advertisement. Copyright 1997, 1998, 1999.

**EXERCISE IV**

Provide a close analysis of the following advertisement by circling and labeling each of the key argumentative terms, using the same abbreviations as in Exercise 1. Then state what you take to be the central conclusions and premises. What criticisms, if any, do you have of the argument?

**ADVERTISEMENT**

This advertisement appeared in various national magazines in 2008.

Fresher Air

Fuel Efficiency

Smooth Ride

Quietness

What will your reason be?

Hybrid Synergy Drive, delivers in more ways than one. Fresher air is the result of the lowest CO<sub>2</sub> emissions in its class\*. Its fuel efficiency is demonstrated by its ability to travel 650 miles on a single tank of petrol, and its electric motor can take you from 0 to 62 mph in just 10.9 seconds. And just when you thought it could not get any better, the Torque on Demand Control uses the high-output electric motors to deliver a responsive, smooth and quiet acceleration. Will you choose it for yourself, for the earth, or both? Hybrid Synergy Drive. What will your reason be?

**HYBRID SYNERGY DRIVE**

[www.HybridSynergyDrive.com](http://www.HybridSynergyDrive.com)

**TODAY TOMORROW TOYOTA**

All figures quoted refer to Toyota Prius with Hybrid Synergy Drive. Official Fuel Consumption Figures for the Prius 1.5 Hybrid in mpg (l/100 km): Urban – 56.5 (3.0), Extra-Urban – 47.3 (4.2), Combined – 65.7 (4.3), CO<sub>2</sub> – 104 g/km. \*Lowest CO<sub>2</sub> emissions than all other vehicles in the upper medium segment, according to SMMT annual CO<sub>2</sub> report 2006 market overview.

**EXERCISE V**

Practice close analysis some more by doing close analyses of:

1. the passage in the Discussion Question at the end of Chapter 1,
2. an editorial, letter to the editor, or advertisement from your local paper,
3. something that you read for another course,
4. a lecture by your professor in another course (or this course!), or
5. a paper by you or by a friend in another course.

**DISCUSSION QUESTIONS**

1. If, as some social critics have maintained, the pervasive nature of television has created generation upon generation of intellectually passive automations, why study close analysis?
2. Television commercials are often arguments in miniature. Recount several recent television commercials and identify the argumentative devices at work.

## DEEP ANALYSIS

*Arguments in everyday life rarely occur in isolation. They usually come in the middle of much verbiage that is not essential to the argument itself. Everyday arguments are also rarely complete. Essential premises are often omitted. Many such omissions are tolerable because we are able to convey a great deal of information indirectly by conversational implication. Nevertheless, to understand and evaluate an argument, it is necessary to isolate the argument from extraneous surroundings, to make explicit unstated parts of the argument, and to arrange them in a systematic order. This reconstruction puts us in a better position to decide how good the argument really is. This chapter will develop methods for reconstructing arguments so that they may be analyzed and assessed in a fair and systematic fashion. These methods will then be illustrated by applying them to an important disagreement that depends on fundamental principles.*

### GETTING DOWN TO BASICS

---

To understand an argument, it is useful to put it into standard form. As we saw in Chapter 3, this is done simply by writing down the numbered premises, drawing a line, adding “.:” followed by the conclusion, and indicating which premises are supposed to be reasons for the conclusion. That is all we write down in standard form, but there is often a lot more in the passage that includes the argument. It is not uncommon for the stated argument to stretch over several pages, whereas the basic argument has only a few premises and a single conclusion.

One reason for this is that people often go off on *tangents*. They start to argue for one claim, but that reminds them of something else, so they talk about that for a while; then they finally return to their original topic. One example occurred during the Republican presidential candidates’ debate on October 9, 2007, when Governor Mitt Romney said,

We’re also going to have to get serious about treating Ahmadinejad [the President of Iran] like the rogue and buffoon that he is. And it was outrageous for the United Nations to invite him to come to this country. It was outrageous for Columbia to invite him to speak at their university. This is a person who denied

the Holocaust, a person who has spoken about genocide, is seeking the means to carry it out. And it is unacceptable to this country to allow that individual to have control of launching a nuclear weapon. And so we will take the action necessary to keep that from happening.<sup>1</sup>

Romney's criticisms of the United Nations and Columbia are not really part of his argument, because they do not support his conclusion that the United States needs to keep nuclear weapons out of the hands of Ahmadinejad.

Such tangents can be completely irrelevant or unnecessary, and they often make it hard to follow the argument. Some people even go off on tangents on purpose to confuse their opponents and hide gaping holes in their arguments. The irrelevant diversion is sometimes called a *red herring* (reportedly after a man who, when pursued by hounds, threw them off his scent by dragging a red herring across his trail). More generally, this maneuver might be called the *trick of excess verbiage*. It violates the conversational rules of Quantity, Relevance, or Manner, which were discussed in Chapter 2.

To focus on the argument itself, we need to look carefully at each particular sentence to determine whether it affects the validity or strength of the argument or the truth of its premises. If we decide that a sentence is not necessary for the argument, then we should not add it when we list the premises and conclusion in standard form. Of course, we have to be careful not to omit anything that would improve the argument, but we also do not want to include too much, because irrelevant material simply makes it more difficult to analyze and evaluate the argument.

Another source of extra material is repetition. Consider Senator John Edwards's response to a question about the Defense of Marriage Act in the Democratic presidential candidates' debate on January 22, 2004:

These are issues that should be left [to the states]. Massachusetts, for example, has just made a decision—the Supreme Court at least has made a decision—that embraces the notion of gay marriage. I think these are decisions the states should have the power to make. And the Defense of Marriage Act, as I understand it—you're right, I wasn't there when it was passed—but as I understand it, would have taken away that power. And I think that's wrong—that power should not be taken away from the states.<sup>2</sup>

Now compare:

These are issues that should be left to the states.

These are decisions that states should have the power to make.

That power should not be taken away from the states.

All three of these sentences say pretty much the same thing, so we do not need them all.

Why do people repeat themselves like this? Sometimes they just forget that they already made the point before, but often repetition accomplishes a goal. Good speakers regularly repeat their main points to remind their audience of what was said earlier. Repetition is subtler when it is used to explain something.

A point can often be clarified by restating it in a new way. Repetition can also function as a kind of assurance, as an expression of confidence, or as an indication of how important a point is. Some writers seem to think that if they say something often enough, people will come to believe it. Whether or not this trick works, if two sentences say equivalent things, there is no need to list both sentences when the argument is put into standard form. Listing the same premise twice will not make the argument any better from a logical point of view.

Sometimes *guarding* terms can also be dropped. If I say, “I think Miranda is at home, so we can probably meet her there,” this argument might be represented in standard form thus:

- (1) I think Miranda is at home.
- ∴ (2) We can probably meet her there. (from 1)

This is misleading. My *thoughts* are not what make us able to meet Miranda at home. My thoughts do not even increase the probability that she is at home or that we can meet her there. It is the *fact* that Miranda is at home that provides a reason for the conclusion. Thus, it is clearer to drop the guarding phrase (“I think”) when putting the argument into standard form. But you have to be careful, for not all guarding phrases can be dropped. When I say “We can *probably* meet her there,” I might not want to say simply, “We *can* meet her there.” After all, even if she is there now, we might not be able to get there before she leaves. Then to drop “probably” from my conclusion would distort what I meant to say and would make my argument more questionable, so you should not drop that guarding term if you want to understand my argument charitably and accurately.

Here’s another example: If a friend says that you ought to buckle your seat belt because you could have an accident, it would distort her argument to drop the guarding term (“could”), because she is not claiming that you definitely will have an accident, or even that you probably will have one. The *chance* of an accident is significant enough to show that you ought to buckle your seat belt, so this guarding term should be kept when the argument is put into standard form.

It is also possible to drop *assuring* terms in some cases. Suppose someone says, “You obviously cannot play golf in Alaska in January, so there’s no point in bringing your clubs.” There is no need to keep the assuring term (“obviously”) in the premise. It might even be misleading, because the issue is whether the premise is true, not whether it is obvious. The argument cannot be refuted by showing that, even though you in fact cannot play golf in Alaska in January, this is not obvious, since there might be indoor golf courses. In contrast, assuring terms *cannot* be dropped in some other cases. For example, if someone argues, “We know that poverty causes crime, because many studies have shown that it does,” then the assuring terms (“We know that . . .” and “studies have shown that . . .”) cannot be dropped without turning the argument into an empty shell: “Poverty causes crime, because it does.” The point of this argument is to cite the sources of

our knowledge (“studies”) and to show that we have knowledge instead of just a hunch. That point is lost if we drop the assuring terms.

Unfortunately, there is no mechanical method for determining when guarding or assuring terms and phrases can be dropped, or whether certain sentences are unnecessary tangents or repetition. We simply have to look closely at what is being said and think hard about what is needed to support the conclusion. It takes great skill, care, and insight to pare an argument down to its essential core without omitting anything that would make it better. And that is the goal: If you want to understand someone’s argument, you should try to make that argument as good as it can be. You should interpret it charitably. Distorting and oversimplifying other people’s arguments might be fun at times and can win points in debates, but it cannot help us understand or learn from other people’s arguments.

#### EXERCISE 1

Put the following arguments into standard form and omit anything that does not affect the validity of the argument or the truth of its premises:

1. Philadelphia is rich in history, but it is not now the capital of the United States, so the U.S. Congress must meet somewhere else.
2. Not everybody whom you invited is going to come to your party. Some of them won’t come. So this room should be big enough.
3. I know that my wife is at home, since I just called her there and spoke to her. We talked about our dinner plans.
4. I’m not sure, but Joseph is probably Jewish. Hence, he is a rabbi if he is a member of the clergy.
5. Some students could not concentrate on the lecture, because they did not eat lunch before class, although I did.
6. The most surprising news of all is that Johnson dropped out of the race because he thought his opponent was better qualified than he was for the office.
7. The liberal candidate is likely to win, since experts agree that more women support him.
8. It seems to me that married people are happier, so marriage must be a good thing, or at least I think so.

#### DISCUSSION QUESTION

In the quotation above (p. 80), is it fair to drop “I think” from the start of Edwards’s sentences “I think these are decisions the states should have the power to make” and “I think that’s wrong—that power should not be taken away from the states”? Why or why not? Is this phrase “I think” used for guarding or assuring or some other purpose in this context? Explain why Edwards adds these words.

## CLARIFYING CRUCIAL TERMS

After the essential premises and conclusion are isolated, we often need to clarify these claims before we can begin our logical analysis. The goal here is not perfect clarity, for there probably is no such thing. It is, however, often necessary to eliminate ambiguity and reduce vagueness before we can give an argument a fair assessment. In particular, it is usually helpful to specify the referents of pronouns, because such references can depend on a context that is changed when the argument is put into standard form. “You are wrong” or “That’s wrong” can be perfectly clear when said in response to a particular claim, but they lose their clarity when they are moved into the conclusion of an argument in standard form. We also often need to specify whether a claim is about all, most, many, or just some of its subject matter. When people say, “Blues music is sad,” do they mean all, most, some, or typical blues music?

Another common problem arises when someone argues like this:

You should just say “No” to drugs, because drugs are dangerous.

What counts as a drug? What about penicillin or aspirin? The speaker might seem to mean “drugs like cocaine,” but “like” them in which respects? Maybe what is meant is “addictive drugs,” but what about alcohol and nicotine (which are often addictive)? You might think that the speaker means “dangerous drugs,” but then the premise becomes empty: “Dangerous drugs are dangerous.” Or maybe the idea is “illegal drugs,” but that seems to assume that the law is correct about what is dangerous. In any case, we cannot begin to evaluate this argument if we do not know the extent of what it claims.

Of course, we should not try to clarify every term in the argument. Even if this were possible, it would make the argument extremely long and boring. Instead, our goal is to clarify anything that seems likely to produce confusion later if it is not cleared up now. As our analysis continues, we can always return and clarify more if the need arises, but it is better to get the most obvious problems out of the way at the start.

Some problems, however, just won’t go away. Don’t get frustrated if you cannot figure out how to clarify a crucial term in someone else’s argument. The fault might lie with the person who gave the argument. Often an argument leaves a crucial term vague or ambiguous, because serious defects in the argument would become apparent if its terms were made more precise. We will discuss such tricks in detail in Chapters 13 and 14. For now, we just need to try our best to understand and clarify the essential terms in the argument.

## DISSECTING THE ARGUMENT

A single sentence often includes several clauses that make separate claims. When this happens, it is usually useful to dissect the sentence into its smallest parts, so that we can investigate each part separately. Because simpler

steps are easier to follow than complex ones, we can understand the argument better when it is broken down. Dissection makes us more likely to notice any flaws in the argument. It also enables us to pinpoint exactly where the argument fails, if it does.

The process of dissecting an argument is a skill that can be learned only by practice. Let's start with a simple example:

Joe won his bet, because all he had to do was eat five pounds of oysters, and he ate nine dozen oysters, which weigh more than five pounds.

The simplest unpacking of this argument yields the following restatement in standard form:

(1) All Joe had to do was eat five pounds of oysters, and he ate nine dozen oysters, which weigh more than five pounds.

∴ (2) Joe won his bet. (from 1)

If we think about the premise of this argument, we see that it actually contains three claims. The argument will be clearer if we separate these claims into independent premises and add a few words for the sake of clarity. The following, then, is a better representation of this argument:

(1) All Joe had to do (to win his bet) was eat five pounds of oysters.

(2) Joe ate nine dozen oysters.

(3) Nine dozen oysters weigh more than five pounds.

∴ (4) Joe won his bet. (from 1–3)

With the premise split up in this way, it becomes obvious that there are three separate ways in which the argument could fail. One possibility is that the first premise is false because Joe had to do more than just eat five pounds of oysters to win his bet: Maybe what he bet was that he could eat five pounds in five minutes. Another possibility is that the second premise is false because Joe did not really eat nine dozen oysters: Maybe he really ate one dozen oysters cut into nine dozen pieces. A final way in which the argument could fail is if the third premise is false because nine dozen oysters do not weigh more than five pounds: Maybe the oysters that Joe ate were very small, or maybe nine dozen oysters weigh more than five pounds only when they are still in their shells, but Joe did not eat the shells. In any case, breaking down complex premises into simpler ones makes it easier to see exactly where the argument goes wrong, if it does. Consequently, we can be more confident that an argument does not go wrong if we do not see any problem in it even after we have broken it down completely.

Although it is a good idea to break down the premises of an argument when this is possible, we have to be careful not to do this in a way that changes the logical structure of the argument. Suppose someone argues like this:

Socialism is doomed to failure because it does not provide the incentives that are needed for a prosperous economy.

The simplest representation of this argument yields the following standard form:

- (1) Socialism does not provide the incentives that are needed for a prosperous economy.
- ∴ (2) Socialism is doomed to failure. (from 1)

It is tempting to break up the first premise into two parts:

- (1) Socialism does not provide incentives.
- (2) Incentives are needed for a prosperous economy.
- ∴ (3) Socialism is doomed to failure. (from 1–2)

In this form, the argument is open to a fatal objection: Socialism *does* provide *some* incentives. Workers often get public recognition and special privileges when they produce a great deal in socialist economies. But this does not refute the original argument. The point of the original argument was not that socialism does not provide any incentives at all, but only that socialism does not provide *enough* incentives or the right *kind* of incentives to create a prosperous economy. This point is lost if we break up the premise in the way suggested. A better attempt is this:

- (1) Socialism does not provide adequate incentives.
- (2) Adequate incentives are needed for a prosperous economy.
- ∴ (3) Socialism is doomed to failure. (from 1–2)

The problem now is to specify when incentives are *adequate*. What kinds of incentives are needed? How much of these incentives? The answer seems to be “enough for a prosperous economy.” But then premise 2 reduces to “Enough incentives for a prosperous economy are needed for a prosperous economy.” This is too empty to be useful. Thus, we are led back to something like the original premise:

- (1) Socialism does not provide enough incentives for a prosperous economy.
- ∴ (2) Socialism is doomed to failure. (from 1)

In this case, we cannot break the premise into parts without distorting the point.

## ARRANGING SUBARGUMENTS

When the premises of an argument are dissected, it often becomes clear that some of these premises are intended as reasons for others. The premises then

form a chain of simpler arguments that culminate in the ultimate conclusion, but only after some intermediate steps. Consider this argument:

There's no way I can finish my paper before the 9 o'clock show, since I have to do the reading first, so I won't even start writing until at least 9 o'clock.

It might seem tempting to put this argument into standard form as:

- (1) I have to do the reading first.
- (2) I won't even start writing until at least 9 o'clock.
- ∴(3) I can't finish my paper before the 9 o'clock show. (from 1–2)

This reformulation does include all three parts of the original argument, but it fails to indicate the correct role for each part. The two argument markers in the original argument indicate that there are really *two* conclusions. The word "since" indicates that what precedes it is a conclusion, and the word "so" indicates that what follows it is also a conclusion. We cannot represent this as a single argument in standard form, because each argument in standard form can have only one conclusion. Thus, the original sentence must have included two arguments. The relationship between these arguments should be clear: The conclusion of the first argument functions as a premise or reason in the second argument. To represent this, we let the two arguments form a chain. This is the first argument:

- (1) I have to do the reading first.
- ∴(2) I won't even start writing until at least 9 o'clock. (from 1)

This is the second argument:

- (2) I won't even start writing until at least 9 o'clock.
- ∴(3) I can't finish my paper before the 9 o'clock show. (from 2)

If we want to, we can then write these two arguments in a chain like this:

- (1) I have to do the reading first.
- ∴(2) I won't even start writing until at least 9 o'clock. (from 1)
- ∴(3) I can't finish my paper before the 9 o'clock show. (from 2)

This chain of reasoning can also be diagrammed like this:



The arrows indicate which claims are supposed to provide reasons for which other claims. Because these premises and arrows all fall on a single line, it is natural to call this structure *linear*.

Although it is often illuminating to break an argument into stages and arrange them in a linear series, this can be misleading if done incorrectly. For example, the first sentences of Kyl's speech cited in Chapter 4 read as follows:

Mr. Speaker, I oppose this measure. I oppose it first because it is expensive. I further oppose it because it is untimely.

If we try to force this into a simple line, we might get this:

- (1) This measure is expensive.
- ∴(2) This measure is untimely. (from 1)
- ∴(3) I oppose this measure. (from 2)

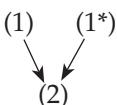
This reconstruction suggests that the measure's being expensive is what makes it untimely. That might be true (say, during a temporary budget crisis), but it is not what Kyl actually says. Instead, Kyl is giving two separate reasons for the same conclusion. First,

- (1) This measure is expensive.
- ∴(2) I oppose this measure. (from 1)

Second,

- (1\*) This measure is untimely.
- ∴(2) I oppose this measure. (from 1\*)

The structure of this argument can now be diagrammed as a branching tree:



The two arrows indicate that there are two separate reasons for the conclusion. Because this structure resembles the way branches split off from the trunk of a tree, we can describe this structure as *branching*. We have to be careful not to confuse branching arguments like this with linear chains of arguments that do not branch.

We also need to distinguish this branching structure from cases where several premises work together to support a single conclusion. Consider this argument:

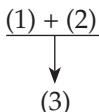
My keys must be either at home or at the office. They can't be at the office, because I looked for them there. So they must be at home.

With some clarifications, we can put this argument in standard form:

- (1) My keys are either at my home or at my office.
- (2) My keys are not at my office.
- ∴(3) My keys are at my home. (from 1–2)

Although this argument has two premises, it does not give two separate reasons for its conclusion. Neither premise by itself, without the other, is enough to give us any reason to believe the conclusion: “My keys are either at my home or at my office” alone is not enough to support “My keys are at my home,” and “My keys are not at my office” alone is also not enough to support “My keys are at my home.” The premises work only when they work together. Thus, it would be misleading to diagram this argument in the same way as Kyl’s argument.

Instead, we need to indicate that the premises work together. Here’s a simple way:



The symbol “+” with a single arrow indicates that the two premises work together to provide a single reason for the conclusion. The line under the premises that are joined together makes it clear that those are the premises that lead to the conclusion at the end of the arrow. If three or more premises provided a single reason, then we could simply add to the list—(1) + (2) + (3), and so on—then draw a line under the premises to show which ones work together. Because these premises work jointly rather than separately, we can call this structure *joint*.

The argument that we are diagramming included one part that we have not incorporated yet:

They can’t be at the office, because I looked for them there.

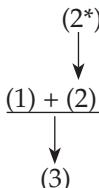
The standard form is this:

(2\*) I looked for my keys at my office.  
      ∴ (2) My keys can’t be at my office. (from 2\*)

By itself, this argument has this diagram:



Since the conclusion of this background argument is a premise in the other part of the argument, we can put the diagrams together like this:



The fact that the arrow goes from (2\*) to (2) but not to (1) indicates that this background argument supports premise (2), but not the other premise. In cases like this, you need to be careful where you draw your arrows.

Argument structures can get very complex, but we can diagram most arguments by connecting the simple forms that we illustrated. Begin by identifying

the premises and conclusions. Give each different claim a different number. When two premises work together to support a single conclusion, put a “+” between the premises and a line under them connected to a single arrow that points to the conclusion. When two or more premises (or sets of premises) provide separate reasons for a conclusion, draw separate arrows from each reason to the conclusion. When a conclusion of one argument is a premise in another, put it in the middle of a chain. The complete diagram together will then show how the parts of the argument fit together and form a complex whole.

### EXERCISE II

Put the following arguments into standard form. Break up the premises and form chains of arguments wherever this can be done without distorting the argument. Then diagram the argument.

1. I know that Pat can't be a father, because she is not a male. So she can't be a grandfather either.
2. Either Jack is a fool or Mary is a crook, because she ended up with all of his money.
3. Our team can't win this Saturday, both because they are not going to play, and because they are no good, so they wouldn't win even if they did play.
4. Mercury is known to be the only metal that is liquid at room temperature, so a pound of mercury would be liquid in this room, which is at room temperature, and it would also conduct electricity, since all metals do. Therefore, some liquids do conduct electricity.
5. Since he won the lottery, he's rich and lucky, so he'll probably do well in the stock market, too, unless his luck runs out.
6. Joe is not a freshman, since he lives in a fraternity, and freshmen are not allowed to live in fraternities. He also can't be a senior, since he has not declared a major, and every senior has declared a major. And he can't be a junior, because I never met him before today, and I would have met him before now if he were a junior. So Joe must be a sophomore.
7. Since many newly emerging nations do not have the capital resources necessary for sustained growth, they will continue to need help from industrial nations to avoid mass starvation.

### EXERCISE III

In “A Piece of ‘God’s Handiwork’” (Exercise II in Chapter 4), Robert Redford argues that the Bureau of Land Management (BLM) should not allow Conoco to drill for oil in Utah’s Grand Staircase-Escalante National Monument. The following passage is a crucial part where Redford answers an objection. Arrange its subarguments in standard form so as to reveal the structure of his argument. Then diagram the overall argument.

*(continued)*

The BLM says its hands are tied. Why? Because these lands were set aside subject to “valid existing rights,” and Conoco has a lease that gives it the right to drill. Sure Conoco has a lease—more than one, in fact—but those leases were originally issued without sufficient environmental study or public input. As a result, none of them conveyed a valid right to drill. What’s more, in deciding to issue a permit to drill now, the BLM did not conduct a full analysis of the environmental impacts of drilling in these incomparable lands, but instead determined there would be no significant environmental harm on the basis of an abbreviated review that didn’t even look at drilling on the other federal leases. Sounds like Washington double-speak to me.<sup>3</sup>

#### EXERCISE IV

During the Republican candidates’ debate on October 9, 2007, Chris Matthews asked Senator John McCain, “... Do you believe that Congress has to authorize a strategic attack, not an attack on—during hot pursuit, but a strategic attack on weaponry in Iran—do you need congressional approval as commander and chief?” Read McCain’s response, then arrange its subarguments in standard form so as to reveal the structure of his argument. Then diagram the overall argument.

McCain: We’re dealing, of course, with hypotheticals. If the situation is that it requires immediate action to ensure the security of the United States of America, that’s what you take your oath to do, when you’re inaugurated as president of the United States. If it’s a long series of build-ups, where the threat becomes greater and greater, of course you want to go to Congress; of course you want to get approval, if this is an imminent threat to the security of the United States of America. So it obviously depends on the scenario. But I would, at minimum, consult with the leaders of Congress because there may come a time when you need the approval of Congress. And I believe that this is a possibility that is, maybe, closer to reality than we are discussing tonight.<sup>4</sup>

### SOME STANDARDS FOR EVALUATING ARGUMENTS

After identifying the explicit premises and conclusion and then placing them all into a unified structure, the next step is to look for missing parts. Arguments in everyday life are rarely completely explicit. They usually depend on unstated assumptions that are taken for granted by those in the conversation. We need to bring out those implicit elements in order to complete the argument and assess it fully.

This step raises a crucial question: When is it legitimate to add premises that the arguer did not state openly? It would be unfair to criticize an

argument for assuming something that it does not really need to assume. Nonetheless, if an argument does need to meet certain standards in order to support its conclusion, then it is legitimate to add premises that really are necessary for the argument to meet those standards. Thus, in order to determine which assumptions we can fairly ascribe to an argument, we first need to determine precisely which standards that argument needs to meet in order to succeed or be good.

Evaluating arguments is a complex business. In fact, this entire book is aimed primarily at developing procedures for doing so. We will find that different standards apply to different arguments. There are, however, certain basic terms used in evaluating many arguments that we can introduce briefly now. They are validity, truth, and soundness. Here they will be introduced informally. Later (in Chapters 6 and 7) they will be examined with more rigor.

## VALIDITY

In some good arguments, the conclusion is said to follow from the premises. However, this commonsense notion of *following from* is hard to pin down precisely. The conclusion follows from the premises only when the content of the conclusion is related appropriately to the content of the premises, but which relations count as appropriate?

To avoid this difficult question, most logicians instead discuss whether an argument is *valid*. Calling something “valid” can mean a variety of things, but in this context validity is a technical notion. Here “valid” does not mean “good,” and “invalid” does not mean “bad.” This will be our definition of validity:

An argument is *valid* if and only if it is not possible that all of its premises are true and its conclusion false.

Alternatively, one could say that its conclusion *must* be true if its premises are all true (or, again, that at least one of its premises *must* be false if its conclusion is false). The point is that a certain combination—true premises and a false conclusion—is ruled out as impossible.

The following argument passes this test for validity:

- (1) All senators are paid.
  - (2) Sam is a senator.
- 
- ∴ (3) Sam is paid. (from 1–2)

Clearly, if the two premises are both true, there is no way for the conclusion to fail to be true. To see this, just try to tell a coherent story in which every single senator is paid and Sam is a senator, but Sam is not paid. You can’t do it.

Contrast this example with a different argument:

- (1) All senators are paid.
- (2) Sam is paid.
- ∴(3) Sam is a senator. (from 1–2)

Here the premises and the conclusion are all in fact true, let's assume, but that is still not enough to make the argument valid, because validity concerns what is possible or impossible, not what happens to be true. This conclusion *could* be false even when the premises are true, for Sam *could* leave the Senate but still be paid for some other job, such as lobbyist. That possibility shows that this argument is invalid.

Another very common form of argument is called *modus ponens*:

- (1) If it is snowing, then the roads are slippery.
- (2) It is snowing.
- ∴(3) The roads are slippery. (from 1–2)

This argument is valid, because it is not possible for its premises to be true when its conclusion is false. We can show that by assuming that the conclusion is false and then reasoning backwards. Imagine that the roads are not slippery. Then there are two possibilities. Either it is snowing or it is not snowing. If it is not snowing, then the second premise is false. If it is snowing, then the first premise must be false, since we are supposing that it is snowing and that the roads are not slippery. Thus, at least one premise has to be false when the conclusion is false. Hence, this argument is valid.

This argument might seem similar to another:

- (1) If it is snowing, then the roads are slippery.
- (2) It is not snowing.
- ∴(3) The roads are not slippery. (from 1–2)

This argument is clearly invalid, because there are several ways for its premises to be true when its conclusion is false. It might have just stopped snowing or ice might make the roads slippery. Then the roads are slippery, so the conclusion is false, even if both premises are true.

Yet another form of argument is often called *process of elimination*:

- (1) Either Joe or Jack or Jim or Jerry committed the murder.
- (2) Joe didn't do it.
- (3) Jack didn't do it.
- (4) Jim didn't do it.
- ∴(5) Jerry committed the murder. (from 1–4)

The first premise asserts that at least one of these four suspects is guilty. That couldn't be true if all of the other premises were true and the conclusion were false, because that combination would exclude all four of these suspects. So this argument is valid.

Now compare this argument:

- (1) Either Joe or Jack or Jim or Jerry committed the murder.
  - (2) Joe did it.
- ∴ (3) Jerry did not commit the murder. (from 1–2)

To show that this argument is invalid, all we have to do is explain how the premises could be true and the conclusion false. Here's how: Joe and Jerry did it together. In that case, Jerry did it, so the conclusion is false; Joe also did it, so the second premise is true; and the first premise is true, because it says that at least one of these four suspects did it, and that is true when more than one of the suspects did it. That possibility of complicity, thus, makes this argument invalid.

We will explore many more forms of argument in Chapters 6 and 7. The goal for now is just to get a feel for how to determine validity. In all of these examples, an argument is said to be *valid* if and only if there is no possible situation in which its premises are true and its conclusion is false. You need to figure out whether there could be any situation like this in order to determine whether an argument is valid. If so, the argument is invalid. If not, it is valid.

This definition shows why validity is a valuable feature for an argument to possess: There can be no valid argument that leads one from true premises to a false conclusion. This should square with your commonsense ideas about reasoning. If you reason well, you should not be led from truth into error.

What are known as *deductive* arguments are put forward as meeting this standard of validity, so validity is one criterion for a good deductive argument. Other arguments—so-called *inductive* arguments—are not presented as meeting this standard. Roughly, an inductive argument is presented as providing strong support for its conclusion. The standards for evaluating inductive arguments will be examined in Chapters 8–10. For now, we will concentrate on deductive arguments.

## TRUTH

Although a deductive argument must be valid in order to be a good argument, validity is not enough. One reason is that an argument can be valid even when some (or all) of the statements it contains are false. For example:

- (1) No fathers are female.
  - (2) Sam is a father.
- ∴ (3) Sam is not female. (from 1–2)

Suppose that Sam has no children or that Sam is female, so premise 2 is false. That would be a serious defect in this argument. Nonetheless, this argument satisfies our definition of validity: If the premises were true, then the conclusion could not be false. There is no way that Sam could be female if Sam is a

father and no fathers are female. This example makes it obvious that validity is not the same as truth. It also makes it obvious that another requirement of a good argument is that *all of its premises must be true*.

## SOUNDNESS

We thus make at least two demands of a deductive argument:

1. The argument must be valid.
2. The premises must be true.

When an argument meets both of these standards, it is said to be *sound*. If it fails to meet either one or the other, then it is *unsound*. Thus, an argument is unsound if it is invalid, and it is also unsound if at least one of its premises is false.

	ALL PREMISES TRUE	AT LEAST ONE FALSE PREMISE
Valid	Sound	Unsound
Invalid	Unsound	Unsound

Soundness has one great benefit: A sound argument must have a true conclusion. We know this because its premises are true and, since it is valid, it is not possible that its premises are true and its conclusion is false. This is why people who seek truth want sound arguments, not merely valid arguments.

## EXERCISE V

Indicate whether each of the following arguments is valid and whether it is sound. Explain your answers where necessary.

1. Most professors agree that they are paid too little, so they are.
2. David Letterman is over four feet tall, so he is over two feet tall.
3. Lee can't run a company right, because he can't do anything right.
4. Barack Obama is smart and good-looking, so he is smart.
5. Barack Obama is either a Democrat or a Republican, so he is a Democrat.
6. Since Jimmy Carter was president, he must have won an election.
7. Since Gerald Ford was president, he must have won an election.
8. Pat is either a mother or a father. If Pat is a mother, then she is a parent. If Pat is a father, he is a parent. So, either way, Pat is a parent. (Assume that this conclusion is true.)
9. People who live in the Carolinas live in either North Carolina or South Carolina. Hillary Clinton does not live in North Carolina or South Carolina. Hence, she does not live in the Carolinas.
10. If all of Illinois were in Canada, then Chicago would be in Canada. But Chicago is not in Canada. Therefore, not all of Illinois is in Canada.

11. If George lives in Crawford, then George lives in Texas. If George lives in Texas, then George lives in the United States. Hence, if George lives in Crawford, he lives in the United States.
12. There can't be a largest six-digit number, because six-digit numbers are numbers, and there is no largest number.

### EXERCISE VI

Assume that the following sentences are either true (T) or false (F) as indicated.

- All my children are teenagers. (T)  
All teenagers are students. (T)  
All teenagers are my children. (F)  
All my children are students. (T)

Using these assigned values, label each of the following arguments as (a) either valid or invalid, and (b) either sound or unsound.

1. All my children are teenagers.  
All teenagers are students.  
∴ All my children are students.
2. All my children are students.  
All teenagers are students.  
∴ All my children are teenagers.
3. All teenagers are my children.  
All my children are students.  
∴ All teenagers are students.
4. All teenagers are students.  
All my children are students.  
∴ All my children are students.

### EXERCISE VII

Indicate whether each of the following sentences is true. For those that are true, explain why they are true. For those that are false, show why they are false by giving a counterexample.

1. Every argument with a false conclusion is invalid.
2. Every argument with a false premise is invalid.
3. Every argument with a false premise and a false conclusion is invalid.
4. Every argument with a false premise and a true conclusion is invalid.

(continued)

5. Every argument with true premises and a false conclusion is invalid.
6. Every argument with a true conclusion is sound.
7. Every argument with a false conclusion is unsound.

## SUPPRESSED PREMISES

Now that we understand validity and soundness, we can use those standards to determine which assumptions can fairly be added to deductive arguments in order to complete them. If some extra premise is needed in order for a deductive argument to be valid or sound, then that argument needs that assumption in order to succeed as a deductive argument. That makes it legitimate to add that extra premise to the argument even though the person who gave that argument omitted that premise. The arguer did not openly state the extra premise, but he did assume it.

For example, if we are told that Chester Arthur was a president of the United States, we have a right to conclude a great many things about him—for example, that at the time he was president, he was a live human being. Appeals to facts of this kind lie behind the following argument:

Benjamin Franklin could not have been our second president, because he died before the second election was held.

This argument obviously turns on a question of fact: Did Franklin die before the second presidential election was held? (He did.) The argument would not be sound if this explicit premise were not true. But the argument also depends on a more general principle that ties the premise and conclusion together:

The dead cannot be president.

This new premise is needed to make the argument valid in the technical sense.

This new premise is also needed to explain *why* the premise supports the conclusion. You could have made the original argument valid simply by adding this:

If Franklin died before the second election was held, then he could not have been our second president.

Indeed, you can always make an argument valid simply by adding a conditional whose antecedent is the premises and whose consequent is the conclusion. However, this trick is often not illuminating; it does not reveal how the argument works. In our example, there is nothing special about Franklin, so it is misleading to add a conditional that mentions Franklin in particular. In contrast, when we add the general principle, “The dead cannot be president,” this new premise not only makes the argument valid but also helps us understand how the conclusion is supposed to follow from the premise.

Traditionally, logicians have called premises that are not stated but are needed (to make the argument valid and explain how it works) *suppressed premises*. An argument depending on suppressed premises is called an *enthymeme* and is said to be *enthymematic*. If we look at arguments that occur in daily life, we discover that they are, almost without exception, enthymematic. Therefore, to trace the pathway between premises and conclusion, it is usually necessary to fill in these suppressed premises that serve as links between the stated premises and the conclusion.

## CONTINGENT FACTS

Suppressed premises come in several varieties. They often concern facts or conventions that might have been otherwise—that are contingent rather than necessary. Our example assumed that the dead are not eligible for the presidency, but we can imagine a society in which the deceased are elected to public office as an honor (something like posthumous induction into the Baseball Hall of Fame). Our national government is not like that, however, and this is something that most Americans know. This makes it odd to come right out and say that the deceased cannot hold public office. In most settings, this would involve a violation of the conversational rule of Quantity, because it says more than needs to be said.

Even though it would be odd to state it, this fact plays a central role in the argument. To assert the conclusion without believing the suppressed premise would involve a violation of the conversational rule of Quality, because the speaker would not have adequate reasons for the conclusion. Furthermore, if this suppressed premise were not believed to be true, then to give the explicit premise as a reason for the conclusion would violate the conversational rule of Relevance (just as it would be irrelevant to point out that Babe Ruth is dead when someone asks whether he is in the Baseball Hall of Fame). For these reasons, anyone who gives the original argument conversationally implies a commitment to the suppressed premise.

Suppressed premises are not always so obvious. A somewhat more complicated example is this:

Arnold Schwarzenegger cannot become president of the United States, because he was born in Austria.

Why should being from Austria disqualify someone from being president? It seems odd that the Founding Fathers should have something against that particular part of the world. The answer is that the argument depends on a more general suppressed premise:

Only a natural-born U.S. citizen may become president of the United States.

It is this provision of the U.S. Constitution that lies at the heart of the argument. Knowing this provision is, of course, a more specialized piece of knowledge than knowing that you have to be alive to be president. For this

reason, more people will see the force of the first argument (about Franklin) than the second (about Schwarzenegger). The second argument assumes an audience with more specialized knowledge.

The argument still has to draw a connection between being born in Austria and being a natural-born U.S. citizen. So it turns out that the argument has three stages:

- (1) Schwarzenegger was born in Austria.
- (2) Austria has never been part of the United States.
- . . . (3) Schwarzenegger was born outside of the United States. (from 1–2)
- (4) Anyone who was born outside of the United States is not a natural-born U.S. citizen.
- . . . (5) Schwarzenegger is not a natural-born U.S. citizen. (from 3–4)
- (6) Only a natural-born U.S. citizen may become president of the United States.
- . . . (7) Schwarzenegger cannot become president of the United States. (from 5–6)

With the addition of suppressed premises (2), (4), and (6), the argument is technically valid, for, if (1)–(2) are true, (3) must be true; if (3)–(4) are true, (5) must also be true; and if (5)–(6) are true, then (7) must be true.

The argument is still not sound, however, because some of the suppressed premises that were added are not true. In particular, there is an exception to the suppressed premise about who is a natural-born U.S. citizen. This exception is well known to U.S. citizens who live overseas. People who were born in Austria are U.S. citizens if their parents were U.S. citizens. They also seem to count as natural-born citizens, since they are not naturalized. This is not completely settled, but it does not matter here, as Arnold Schwarzenegger's parents were not U.S. citizens when he was born. Thus, the second stage of the above argument can be reformulated as follows:

- (3) Schwarzenegger was born outside of the United States.
- (4\*) Schwarzenegger's parents were not U.S. citizens when he was born.
- (4\*\*) Anyone who was born outside of the United States and whose parents were not U.S. citizens at the time is not a natural-born U.S. citizen.
- . . . (5) Schwarzenegger is not a natural-born U.S. citizen. (from 3, 4\*, and 4\*\*)

This much of the argument is now sound.

An argument with a single premise has grown to include three stages with at least four suppressed premises. Some of the added premises are obvious, but others are less well known, so we cannot assume that the person who gave the original argument had the more complete argument in mind. Many people would be convinced by the original argument even without all these added complexities. Nonetheless, the many suppressed premises are necessary to make the argument sound. Seeing this brings out the assumptions that must

be true for the conclusion to follow from the premises. This process of making everything explicit enables us to assess these background assumptions directly.

### EXERCISE VIII

There is one obscure exception to the premise that only a natural-born citizen may become president of the United States. The Constitution does allow a person who is not a natural-born citizen to become president if he or she was “a citizen of the United States at the time of the adoption of this Constitution.” This exception is said to have been added to allow Alexander Hamilton to run for president, but it obviously does not apply to Schwarzenegger or to anyone else alive today. Nonetheless, this exception keeps the argument from being sound in its present form. Reformulate the final stage of the argument to make it sound.

## LINGUISTIC PRINCIPLES

Often an argument is valid, but it is still not clear *why* it is valid. It is not clear *how* the conclusion follows from the premises. Arguments are like pathways between premises and conclusions, and some of these pathways are more complicated than others. Yet even the simplest arguments reveal hidden complexities when examined closely. For example, there is no question that the following argument is valid:

- (1) Harriet is in New York with her son.  
∴(2) Harriet’s son is in New York.

It is not possible for the premise to be true and the conclusion false. If asked why this conclusion follows from the premise, it would be natural to reply that:

You cannot be someplace with somebody unless that person is there, too.

This is not something we usually spell out, but it is the principle that takes us from the premise to the conclusion.

One thing to notice about this principle is that it is quite general—that is, it does not depend on any special features of the people or places involved. It is also true that if Benjamin is in St. Louis with his daughter, then Benjamin’s daughter is in St. Louis. Although the references have changed, the general pattern that lies behind this inference will seem obvious to anyone who understands the words used to formulate it. For this reason, principles of this kind are basically *linguistic* in character.

If we look at arguments as they occur in everyday life, we will discover that almost all of them turn on unstated linguistic principles. To cite just one more example: Alice is taller than her husband, so there is at least one woman who is taller than at least one man. This inference relies on the principles that

husbands are men and wives are women. We do not usually state these linguistic principles, for to do so will often violate the rule of Quantity. (Try to imagine a context in which you would come right out and say, "Husbands, you know, are men." Unless you were speaking to someone just learning the language, this would be a peculiar remark.) Nonetheless, even if it would usually be peculiar to come right out and state such linguistic principles, our arguments still typically presuppose them. This observation reveals yet another way in which our daily use of language moves within a rich, though largely unnoticed, framework of linguistic rules, as we emphasized in Chapter 2.

### EVALUATIVE SUPPRESSED PREMISES

We have examined two kinds of suppressed premises, factual and linguistic. Many arguments also contain unstated evaluative premises. As we saw in Chapter 3, evaluation comes in many kinds. The following argument involves moral evaluation:

It is immoral to buy pornography, because pornography leads to violence toward women.

This argument clearly relies on the moral principle that it is immoral to buy anything that leads to violence toward women. A different example contains religious premises:

You shouldn't take the name of the Lord in vain, for this shows disrespect.

The suppressed premise here is that you should not do anything that shows disrespect (to the Lord). One more example is about economics:

It is unwise to invest all of your money in one stock, since this increases the risk that you will lose everything.

The suppressed premise here is that it is unwise to increase the risk that you will lose everything. More examples could be given, but the point should be clear. Most arguments depend on unstated assumptions, and many of these assumptions are evaluative in one way or another.

### USES AND ABUSES OF SUPPRESSED PREMISES

Talk about *suppressed* premises may bring to mind suppressing a rebellion or an ugly thought, and using *hidden* premises may sound somewhat sneaky. However, the way we are using them, these expressions do not carry such negative connotations. A suppressed or hidden premise is simply an *unstated* premise. It is often legitimate to leave premises unstated. It is legitimate if (1) those who are given the argument can easily supply these unstated premises for themselves, and (2) the unstated premises are not themselves controversial. If done properly, the suppression of premises can add greatly to the efficiency of language. Indeed, without the judicious suppression of obvious premises, many arguments would become too cumbersome to be effective.

On the other hand, suppressed premises can also be used improperly. People sometimes suppress questionable assumptions so that their opponents will not notice where an argument goes astray. For example, when election debates turn to the topic of crime, we often hear arguments like this:

My opponent is opposed to the death penalty, so he must be soft on crime.

The response sometimes sounds like this:

Since my opponent continues to support the death penalty, he must not have read the most recent studies, which show that the death penalty does not deter crime.

The first argument assumes that anyone who is opposed to the death penalty is soft on crime, and the second argument assumes that anyone who read the studies in question would be convinced by them and would turn against the death penalty. Both of these assumptions are questionable, and the questions they raise are central to the debate. If we want to understand these issues and address them directly, we have to bring out these suppressed premises openly.

### EXERCISE IX

The following arguments depend for their validity on suppressed premises of various kinds. For each of them, list enough suppressed premises to make the argument valid and also to show why it is valid. This might require several suppressed premises of various kinds.

**EXAMPLE:** Carol has no sisters, because all her siblings are brothers.

**SUPPRESSED PREMISES:** A sister would be a sibling.  
A brother is not a sister.

1. Britney Spears is under age thirty-five. Therefore, she cannot run for president of the United States.
2. Nixon couldn't have been president in 1950 because he was still in the Senate.
3. 81 is not a prime number, because 81 is divisible by 3.
4. There's no patient named Rupert here; we have only female patients.
5. Columbus did not discover the New World because the Vikings explored Newfoundland centuries earlier.
6. There must not be any survivors, since they would have been found by now.
7. Lincoln could not have met Washington, because Washington was dead before Lincoln was born.
8. Philadelphia cannot play Los Angeles in the World Series, since they are both in the National League.

9. Mildred must be over forty-three, since she has a daughter who is thirty-six years old.
10. He cannot be a grandfather because he never had children.
11. That's not modern poetry; you can understand it.
12. Harold can't play in the Super Bowl, because he broke his leg.
13. Shaquille must be a basketball player, since he is so tall.
14. Dan is either stupid or very cunning, so he must be stupid.
15. Susan refuses to work on Sundays, which shows that she is lazy and inflexible.
16. Jim told me that Mary is a professor, so she can't be a student, since professors must already have degrees.
17. This burglar alarm won't work unless we are lucky or the burglar uses the front door, so we can't count on it.
18. His natural talents were not enough; he still lost the match because he had not practiced sufficiently.

## THE METHOD OF RECONSTRUCTION

---

We can summarize the discussion so far by listing the steps to be taken in reconstructing an argument. The first two steps were discussed in Chapters 4 and 3, respectively.

1. Do a *close analysis* of the passage containing the argument.
2. List all explicit premises and the conclusion in *standard form*.
3. *Clarify* the premises and the conclusion where necessary.
4. *Break up* the premises and the conclusion into smaller parts where this is possible.
5. *Arrange* the parts of the argument into a chain or tree of subarguments where this is possible.
6. Assess each argument and subargument for *validity*.<sup>5</sup>
7. If any argument or subargument is not valid, or if it is not clear why it is valid, add *suppressed premises* that will show how to get from the premises to the conclusion.
8. Assess the *truth* of the premises.

Remember that the goal of reconstruction is not just technical validity but is, instead, to understand why and how the conclusion is supposed to follow from the premises.

After reconstructing the argument, it is often helpful to add some indication of its structure. This can be done by numbering the premises and then, after

each conclusion, listing the premises from which that conclusion follows. (We did this in our examples.) The argument's structure can also be shown by a diagram like those discussed above. Either way, we need to make it clear exactly how the separate parts of the argument are supposed to fit together.

This method is not intended to be mechanical. Each step requires care and intelligence. As a result, a given argument can be reconstructed in various ways with varying degrees of illumination and insight. The goal of this method is to reveal as much of the structure of an argument as possible and to learn from it as much as you can. Different reconstructions approach this goal more or less closely.

The whole process is more complex than our discussion thus far has suggested. This is especially clear in the last three steps of reconstruction, which must be carried out simultaneously. In deciding whether an argument is acceptable, we try to find a set of true suppressed premises that, if added to the stated premises, yields a sound argument for the conclusion. Two problems typically arise when we make this effort:

1. We find a set of premises strong enough to support the conclusion, but at least one of these premises is false.
2. We modify the premises to avoid falsehood, but the conclusion no longer follows from them.

The reconstruction of an argument typically involves shifting back and forth between the demand for a valid argument and the demand for true premises. Eventually, either we show the argument to be sound or we abandon the effort. In the latter case, we conclude that the argument in question has no sound reconstruction. It is still possible that *we* were at fault in not finding a reconstruction that showed the argument to be sound. Perhaps we did not show enough ingenuity in searching for a suppressed premise that would do the trick. There is, in fact, no purely formal or mechanical way of dealing with this problem. A person presenting an argument may reasonably leave out steps, provided that they can easily be filled in by those to whom the argument is addressed. So, in analyzing an argument, we should be charitable, but our charity has limits. After a reasonable search for those suppressed premises that would show the argument to be sound, we should not blame ourselves if we fail to find them. Rather, the blame shifts to the person who formulated the argument for not doing so clearly.

#### EXERCISE X

Reconstruct and diagram the main arguments in:

1. The passages at the end of Chapters 1 and 4.
2. An editorial from your local paper.
3. Your last term paper or a friend's last term paper.

**EXERCISE XI**

Not all arguments are serious or good. The following silly argument comes from a famous scene in *Monty Python and the Holy Grail*. Reconstruct the argument that is supposed to show that the woman is a witch.

**CROWD:** We have found a witch. May we burn her? . . .

**WOMAN:** I'm not a witch! I'm not a witch! . . .

**LEADER:** What makes you think she is a witch?

**MAN #1:** She turned me into a newt!

**LEADER:** A newt?

**MAN #1:** I got better.

**CROWD:** Burn her anyway!

**LEADER:** Quiet! Quiet! There are ways of telling whether she is a witch.

**CROWD:** Are there? What are they? Tell us. Do they hurt?

**LEADER:** Tell me, what do you do with witches?

**CROWD:** Burn them!

**LEADER:** And what do you burn apart from witches?

**MAN #2:** More witches!

**MAN #3:** Wood.

**LEADER:** So, why do witches burn?

**MAN #1:** 'Cause they're made of wood.

**LEADER:** Good! . . . So, how do we tell whether she is made of wood?

**CROWD:** Build a bridge out of her.

**LEADER:** Ah, but can you not also make bridges out of stone?

**CROWD:** Oh yeah.

**LEADER:** Does wood sink in water?

**CROWD:** No, it floats. Throw her into the pond!

**LEADER:** What also floats in water?

**CROWD:** Bread. Apples. Very small rocks. Cider! Great gravy. Cherries. Mud. Churches. Lead.

**ARTHUR:** A duck!

**LEADER:** Exactly. So, logically, —

**MAN #3:** If she weighs the same as a duck, she's made of wood.

**LEADER:** And therefore?

**CROWD:** A witch! . . . A duck. A duck. Here's a duck!

**LEADER:** We shall use my largest scales.

**CROWD:** Burn the witch! (Woman is placed on scales opposite a duck.)

**LEADER:** Remove the supports. (Woman balances duck.)

**CROWD:** A witch!

**WOMAN:** It's a fair cop.

## AN EXAMPLE OF RECONSTRUCTION: CAPITAL PUNISHMENT

We can illustrate the methods of reconstruction by examining the difficult question of the constitutionality of capital punishment. It has been argued that the U.S. Supreme Court should declare the death penalty unconstitutional because it is a cruel and unusual punishment. The explicitly stated argument has the following basic form:

- (1) The death penalty is a cruel and unusual punishment.
- ∴ (2) The death penalty should be declared unconstitutional. (from 1)

This argument plainly depends on two suppressed premises:

SP1: The Constitution prohibits cruel and unusual punishments.

SP2: Anything that the Constitution prohibits should be declared unconstitutional.

These premises and this entire argument refer to the relevant jurisdiction, which is the United States. So the argument, more fully spelled out, looks like this:

- (1) The death penalty is a cruel and unusual punishment.
- (2) SP: The Constitution prohibits cruel and unusual punishments.
- ∴ (3) The Constitution prohibits the death penalty. (from 1–2)
- (4) SP: Anything that the Constitution prohibits should be declared unconstitutional.
- ∴ (5) The death penalty should be declared unconstitutional. (from 3–4)

This reconstruction seems to be a fair representation of the intent of the original argument.

We can now turn to an assessment of this argument. First, the argument is valid: Given the premises, the conclusion does follow. All that remains is to determine the truth of the premises one by one.

Premise 4 seems uncontroversial. This premise is so much an accepted part of our system that no one would challenge it in a courtroom proceeding today.

Premise 2 is clearly true, for the U.S. Constitution does, in fact, prohibit cruel and unusual punishments. Its Eighth Amendment reads, “Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted.” It is not clear, however, just what this prohibition amounts to. In particular, does the punishment have to be *both* cruel and unusual to be prohibited, or is it prohibited whenever it is *either* cruel or unusual? This would make a big difference if cruel punishments were usual, or if some unusual punishments were not cruel. For the moment, let us interpret the language as meaning “both cruel and unusual.”

The first premise—"The death penalty is a cruel and unusual punishment"—obviously forms the heart of the argument. What we would expect, then, is a good supporting argument to be put forward on its behalf. The following argument by Supreme Court Justice Potter Stewart (in *Furman v. Georgia*, 408 U.S. 239 at 309–310 [1972]) was intended to support this claim in particular cases in which the death penalty was imposed for rape and murder:

In the first place, it is clear that these sentences are "cruel" in the sense that they excessively go beyond, not in degree but in kind, the punishments that the state legislatures have determined to be necessary. . . . In the second place, it is equally clear that these sentences are "unusual" in the sense that the penalty of death is infrequently imposed for murder, and that its imposition for rape is extraordinarily rare. But I do not rest my conclusion upon these two propositions alone. These death sentences are cruel and unusual in the same way that being struck by lightning is cruel and unusual. For, of all the people convicted of rapes and murders in 1967 and 1968, many just as reprehensible as these, the petitioners are among a capriciously selected random handful upon whom the sentence of death has in fact been imposed. My concurring brothers [the Justices who agree with Stewart] have demonstrated that, if any basis can be discerned for the selection of these few to be sentenced to die, it is the constitutionally impermissible basis of race.<sup>6</sup>

The first sentence argues that the death penalty is *cruel*. The basic idea is that punishments are cruel if they inflict harms that are much worse than what is necessary for any legitimate and worthwhile purpose. Stewart then seems to accept the state legislatures' view that the death penalty does go far beyond what is necessary. This makes it cruel.

Now let us concentrate on the part of this argument intended to show that the death penalty is an *unusual* punishment. Of course, in civilized nations, the death penalty is reserved for a small range of crimes, but this is hardly the point at issue. The point of the argument is that the death penalty is unusual even for those crimes that are punishable by death, including first-degree murder. Moreover, Stewart claims that, among those convicted of crimes punishable by death, who actually receives a death sentence is determined either capriciously or on the basis of race. The point seems to be that whether a person who is convicted of a capital crime will be given the death penalty depends on the kind of legal aid he or she receives, the prosecutor's willingness to offer a plea bargain, the judge's personality, the beliefs and attitudes of the jury, and many other considerations. At many points in the process, choices that affect the outcome could be based on mere whim or caprice, or even on the race of the defendant or the victim. Why are these factors mentioned? Because, as Stewart says, it is unconstitutional for sentencing to be based on caprice or race.

We can then restate this supporting argument more carefully:

- (1) Very few criminals who were found guilty of crimes that are punishable by death are actually sentenced to death.
- (2) Among those found guilty of crimes punishable by death, who is sentenced to death depends on caprice or race.
- (3) It is unconstitutional for sentencing to depend on caprice or race.
- (4) A punishment is unusual if it is imposed infrequently and on an unconstitutional basis.

∴ (5) The death penalty is an unusual punishment. (from 1–4)

This conclusion is part of the first premise in our original argument. Now we can spread the entire argument out before us:

- (1) An act is cruel if it inflicts harms that are much worse than what is necessary for any legitimate and worthwhile purpose.
  - (2) The death penalty inflicts harms that are much worse than what is necessary for any legitimate and worthwhile purpose.
- ∴ (3) The death penalty is cruel. (from 1–2)
- (4) Very few criminals who were found guilty of crimes that are punishable by death are sentenced to death.
  - (5) Among those found guilty of crimes punishable by death, who is sentenced to death depends on caprice or race.
  - (6) It is unconstitutional for sentencing to depend on caprice or race.
  - (7) A punishment is unusual if it is imposed infrequently and on an unconstitutional basis.
- ∴ (8) The death penalty is an unusual punishment. (from 4–7)
- ∴ (9) The death penalty is both cruel and unusual. (from 3 and 8)
- (10) The Constitution prohibits cruel and unusual punishments.
- ∴ (11) The Constitution prohibits the death penalty. (from 9–10)
- (12) Anything that the Constitution prohibits should be declared unconstitutional.
- ∴ (13) The death penalty should be declared unconstitutional. (from 11–12)

These propositions provide at least the skeleton of an argument with some force. The conclusion does seem to follow from the premises, and the premises themselves seem plausible. We have produced a charitable reconstruction of the argument.

This reconstruction enables us to see precisely how opponents can respond to the argument. Some opponents might deny Premise 2 and claim to the contrary that the death penalty does serve a legitimate purpose, such as retribution, deterrence, or incapacitation. Other opponents might deny Premise 5 and claim that courts do have good reasons for the death sentences that they approve. A more subtle objection denies Premise 7, because

it is not the death penalty itself that is unusual or unfair in the relevant sense when the conditions in premise 7 are met. Instead, the problem is with the present administration of the death penalty. If so, maybe what we need is procedural reform instead of abolishing the death penalty itself. Of course, there are other objections as well as replies to every objection, so the debate goes on.

The point here is only that reconstructing the argument step by step using the method outlined in this chapter makes us able to understand the argument better and to determine more precisely whether and where it is vulnerable or not. This method can thereby help opponents understand each other and deal with their basic disagreements in an intelligent, humane, and civilized way.

### EXERCISE XII

What is the best argument that Justice Stewart could give in support of the premise that the death penalty “excessively go[es] beyond” what is necessary for any legitimate and worthwhile purpose? Is this argument adequate to justify this premise? (For one such argument, see Justice Brennan’s opinion in *Furman v. Georgia*.)

### EXERCISE XIII

To solve a mystery, you need to determine which facts are crucial and then argue from those facts to a solution. Solve the following mysteries and reconstruct your own argument for your solution. These stories come from *Five-Minute Whodunits*, by Stan Smith (New York: Sterling, 1997). The first passage introduces our hero:

Even those acquainted with Thomas P. Stanwick are often struck by his appearance. A lean and lanky young man, he stands six feet two inches tall. His long, thin face is complemented by a full head of brown hair and a droopy mustache. Though not husky in build, he is surprisingly strong and enjoys ruggedly good health. His origins and early life are obscure. He is undeniably well educated, however, for he graduated with high honors from Dartmouth College as a philosophy major.<sup>7</sup>

### MYSTERY 1: A MERE MATTER OF DEDUCTION

Thomas P. Stanwick, the amateur logician, removed a pile of papers from the extra chair and sat down. His friend Inspector Matthew Walker had just returned to his office from the interrogation room, and Stanwick thought he looked unusually weary.

Source: Stanley Smith, “A Mere Matter of Deduction,” from Five-Minute Whodunits. Copyright © 1997 by Stanley Smith. Reprinted with permission of Sterling Publishing Co., Inc., NY, NY.

"I'm glad you dropped by, Tom," said Walker. "We have a difficult case on hand. Several thousand dollars' worth of jewelry was stolen from Hoffman's Jewel Palace yesterday morning. From some clues at the scene and a few handy tips, we have it narrowed down to three suspects: Addington, Burke, and Chatham. We know that at least one of them was involved, and possibly more than one."

"Burke has been suspected in several other cases, hasn't he?" asked Stanwick as he filled his pipe.

"Yes, he has," Walker replied, "but we haven't been able to nail him yet. The other two are small potatoes, so what we really want to know is whether Burke was involved in this one."

"What have you learned about the three of them?"

"Not too much. Addington and Burke were definitely here in the city yesterday. Chatham may not have been. Addington never works alone, and carries a snub-nosed revolver. Chatham always uses an accomplice, and he was seen lurking in the area last week. He also refuses to work with Addington, who he says once set him up."

"Quite a ragamuffin crew!" Stanwick laughed. "Based on what you've said, it's not too hard to deduce whether Burke was involved."

Was Burke involved or not?

## MYSTERY 2: TRIVIA AND SIGNIFICA

"For April, this is starting out to be a pretty quiet month," remarked Inspector Walker as he rummaged in his desk drawer for a cigar.

Thomas P. Stanwick, the amateur logician, finished lighting his pipe and leaned back in his chair, stretching his long legs forward.

"That is indeed unusual," he said. "Spring usually makes some young fancies turn to crime. The change is welcome."

"Not that we police have nothing to do." Walker lit his cigar. "A couple of the youth gangs, the Hawks and the Owls, have been screeching at each other lately. In fact, we heard a rumor that they were planning to fight each other this Wednesday or Thursday, and we're scrambling around trying to find out whether it's true."

"The Hawks all go to Royston North High, don't they?" asked Stanwick.

"That's right. The Owls are the street-smart dropouts who hang out at Joe's Lunch Cafe on Lindhurst. You know that only those who eat at Joe's collect green matchbooks?"

Stanwick blinked and smiled. "I beg your pardon?"

"That's right." Walker picked up a few papers from his desk. "That's the sort of trivia I'm being fed in my reports. Not only that, but everyone at Royston North High wears monogrammed jackets. What else have I got here? Only kids who hang out on Laraby Street fight on weekdays. Laraby is three blocks from Lindhurst. The Hawks go out for pizza three times a week."

(continued)

"Keep going," chuckled Stanwick. "It's wonderful."

"A hog for useless facts, eh? No one who eats at Joe's wears a monogrammed jacket. The Owls elect a new leader every six months, the Hawks every year. Elections! Furthermore, everyone who hangs out on Laraby Street collects green matchbooks. Finally, the older (but not wiser) Owls buy beer at Johnny's Package Store."

Stanwick laughed heartily. "Lewis Carroll," he said, "the author of *Symbolic Logic* and the 'Alice in Wonderland' books, taught logic at Oxford, and he used to construct soriteses, or polysyllogisms [that is, chains of categorical syllogisms], out of material like that. In fact, his were longer and much wilder and more intricate, but of course they were fiction.

"As it is, the information you've cited should ease your worries. Those gangs won't get together to fight until at least Saturday."

How does he know?

## NOTES

<sup>1</sup> CQ Transcripts Wire, Tuesday, October 9, 2007, washingtonpost.com.

<sup>2</sup> LexisNexis™ Academic, Copyright 2004 Federal News Service, Inc.

<sup>3</sup> Robert Redford, "A Piece of 'God's Handiwork,'" *The Washington Post*, November 25, 1997, A19.

<sup>4</sup> CQ Transcripts Wire, Tuesday, October 9, 2007, washingtonpost.com.

<sup>5</sup> We assess inductive arguments for *strength* instead of validity, but here we focus on deductive arguments. Inductive arguments will be examined in Part III, Chapters 8–12.

<sup>6</sup> Supreme Court Justice Potter Stewart, Concurring Opinion, *Furman v. Georgia*, 408 U.S. 239 at 309–310 [1972].

<sup>7</sup> Stan Smith, *Five-Minute Whodunits* (New York: Sterling, 1997).

## HOW TO EVALUATE ARGUMENTS: DEDUCTIVE STANDARDS

After isolating, laying out, and filling in an argument, the next step is to determine whether that uncovered argument is any good. This assessment, like other evaluations, requires standards. There are two main standards for evaluating arguments: the deductive standard of validity and the inductive standard of strength. Part II (which includes Chapters 6 and 7) will investigate the deductive standard of validity. Part III (which includes Chapters 8–12) will then explore the inductive standard of strength.

We already saw in Chapter 5 that an argument is valid in our technical sense if and only if it is not possible that its premises are true and its conclusion false. That standard sounds simple, but it is not so easy to say how to determine whether this combination of truth values is or is not possible in a particular case. Sometimes the validity of an argument can be seen simply by looking at the premises and conclusion viewed as whole propositions. That is the approach of propositional (or sentential) logic, which is the topic of Chapter 6. Another possibility is that the validity of an argument can be seen only by looking inside premises and conclusions to their parts, including their subjects and predicates. That is the approach of categorical (or syllogistic) logic, which is the topic of Chapter 7.

These relatively simple examples of formal logic do not, of course, exhaust the possibilities. There are many more kinds of formal logic. Many arguments remain valid, even though their validity is not captured by either propositional or categorical logic. That creates problems that we will face throughout Chapters 6 and 7. Still, by exploring some simple ways in which arguments can be valid by virtue of their form alone, we can gain greater insight into the nature of validity and, thereby, into the standards for assessing arguments.



## PROPOSITIONAL LOGIC

This chapter begins our investigation of evaluating arguments by means of formal deductive logic. The first part of the chapter will show how the crucial standard of validity, which was introduced in Chapter 5, can be developed rigorously in one area—what is called propositional logic. This branch of logic deals with connectives such as “and” and “or,” which allow us to build up compound propositions from simpler ones. Throughout most of the chapter, the focus will be theoretical rather than immediately practical. It is intended to provide insight into the concept of validity by examining it in an ideal setting. The chapter will close with a discussion of the relationship between the ideal language of symbolic logic and the language we ordinarily speak.

### THE FORMAL ANALYSIS OF ARGUMENTS

---

When we carry out an informal analysis of an argument, we pay close attention to the key words used to present the argument and then ask ourselves whether these key terms have been used properly. So far, we have no exact techniques for answering the question of whether a word is used correctly. We rely, instead, on linguistic instincts that, on the whole, are fairly good.

In a great many cases, people can tell whether an argument marker, such as “therefore,” is used correctly in indicating that one claim follows from another. However, if we go on to ask the average intelligent person *why* one claim follows from the other, he or she will probably have little to say except, perhaps, that it is just obvious. In short, it is often easy to see *that* one claim follows from another, but to explain *why* can be difficult. The purpose of this chapter is to provide such an explanation for some arguments.

This quality of “following from” is elusive, but it is related to the technical notion of validity, which was introduced in Chapter 5. The focus of our attention will be largely on the *concept* of validity. We are not, for the time being at least, interested in whether this or that argument is valid; we want to understand validity itself. To this end, the arguments we will examine are so simple that you will not be able to imagine anyone not understanding them at a glance. Who needs logic to deal with arguments of this kind?

There is, however, good reason for dealing with simple—trivially simple—arguments at the start. The analytic approach to a complex issue is first to break it down into subissues, repeating the process until we reach problems simple enough to be solved. After these simpler problems are solved, we can reverse the process and construct solutions to larger and more complex problems. When done correctly, the *result* of such an analytic process may seem dull and obvious—and it often is. The *discovery* of such a process, in contrast, often demands the insight of genius.

The methods of analysis to be discussed here are *formal* in a specific way. In Chapter 5, we gave the following argument as an example of a valid argument: “All Senators are paid, and Sam is a Senator, so Sam is paid.” The point could have been made just as well with many similar examples: (a) “All Senators are paid, and Sally is a Senator, so Sally is paid.” (b) “All plumbers are paid, and Sally is a plumber, so Sally is paid.” (c) “All plumbers are dirty, and Sally is a plumber, so Sally is dirty.” These arguments are all valid (though not all are sound). Thus, we can change the person we are talking about, the group that we say the person is in, and the property that we ascribe to the person and to the group, all without affecting the validity of the argument at all. That flexibility shows that the validity of this argument does not depend on the particular content of its premises and conclusion. Instead, the validity of this argument results solely from its form. Formal validity of this kind is what formal logics try to capture.

## BASIC PROPOSITIONAL CONNECTIVES

### CONJUNCTION

The first system of formal logic that we will examine concerns propositional (or sentential) connectives. *Propositional connectives* are terms that allow us to build new propositions from old ones, usually combining two or more propositions into a single proposition. For example, given the propositions “John is tall” and “Harry is short,” we can use the term “and” to *conjoin* them, forming a single compound proposition: “John is tall and Harry is short.”

Let us look carefully at the simple word “and” and ask how it functions. “And” is a curious word, for it does not seem to stand for anything, at least in the way in which a proper name (“Churchill”) and a common noun (“dog”) seem to stand for things. Instead of asking what this word stands for, we can ask a different question: What *truth conditions* govern this connective? That is, under what conditions are propositions containing this connective true? To answer this question, we imagine every possible way in which the component propositions can be true or false. Then, for each combination, we decide what truth value to assign to the

entire proposition. This may sound complicated, but an example will make it clear:

John is tall.	Harry is short.	John is tall and Harry is short.
T	T	T
T	F	F
F	T	F
F	F	F

Here the first two columns cover every possibility for the component propositions to be either true or false. The third column states the truth value of the whole proposition for each combination. Clearly, the conjunction of two propositions is true if both of the component propositions are true; otherwise, it is false.

Our reflections have not depended on the particular propositions in our example. We could have been talking about dinosaurs instead of people, and we still would have come to the conclusion that the conjunction of two propositions is true if both propositions are true, but false otherwise. This neglect of the particular content of propositions is what makes our account *formal*.

To reflect the generality of our concerns, we can drop the reference to particular sentences altogether and use variables instead. Just as the lowercase letters “*x*,” “*y*,” and “*z*” can be replaced by any numbers in mathematics, so we can use the lowercase letters “*p*,” “*q*,” “*r*,” “*s*,” and so on as variables that can be replaced by any propositions in logic. We will also use the symbol “&” (called an *ampersand*) for “and.”

Consider the expression “*p & q*.” Is it true or false? There is obviously no answer to this question. This is not because we do not know what “*p*” and “*q*” stand for, for in fact “*p*” and “*q*” do not stand for any proposition at all. Just as “*x + y*” is not any particular number in mathematics, so “*p & q*” is not a proposition. Instead, “*p & q*” is a pattern for a whole series of propositions. To reflect this, we will say that “*p & q*” is a *propositional form*. It is a pattern, or form, for a whole series of propositions, including “John is tall and Harry is short” as well as many other propositions.

To specify precisely which propositions have the form “*p & q*,” we need a little technical terminology. The central idea is that we can pass from a proposition to a propositional form by replacing propositions with propositional variables.

Proposition	Propositional Form
John is tall and Harry is short.	<i>p &amp; q</i>

When we proceed in the opposite direction by uniformly substituting propositions for propositional variables, we get what we will call a *substitution instance* of that propositional form.

Propositional Form	Substitution Instance
<i>p &amp; q</i>	Roses are red and violets are blue.

Thus, “John is tall and Harry is short” and “Roses are red and violets are blue” are both substitution instances of the propositional form “ $p \ \& \ q$ .”

To get clear about these ideas, it is important to notice that “ $p$ ” is also a propositional form, with *every* proposition, including “Roses are red and violets are blue,” among its substitution instances. There is no rule against substituting compound propositions for propositional variables. Perhaps a bit more surprisingly, our definitions allow “Roses are red and roses are red” to be a substitution instance of “ $p \ \& \ q$ .” This example makes sense if you compare it to variables in mathematics. Using only positive integers, how many solutions are there to the equation “ $x + y = 4$ ”? There are three: 3 + 1, 1 + 3, and 2 + 2. The fact that “2 + 2” is a solution to “ $x + y = 4$ ” shows that “2” can be substituted for both “ $x$ ” and “ $y$ ” in the same solution. That’s just like allowing “Roses are red” to be substituted for both “ $p$ ” and “ $q$ ,” so that “Roses are red and roses are red” is a substitution instance of “ $p \ \& \ q$ ” in propositional logic.

In general, then, we get a substitution instance of a propositional form by uniformly replacing the same variable with the same proposition throughout, but different variables do not have to be replaced with different propositions. The rule is this:

Different variables may be replaced with the same proposition, but different propositions may not be replaced with the same variable.

According to this rule:

- “Roses are red and violets are blue” is a substitution instance of “ $p \ \& \ q$ .”
- “Roses are red and violets are blue” is also a substitution instance of “ $p$ .”
- “Roses are red and roses are red” is a substitution instance of “ $p \ \& \ q$ .”
- “Roses are red and roses are red” is a substitution instance of “ $p \ \& \ p$ .”
- “Roses are red and violets are blue” is *not* a substitution instance of “ $p \ \& \ p$ .”
- “Roses are red” is *not* a substitution instance of “ $p \ \& \ p$ .”

We are now in a position to give a perfectly general definition of conjunction with the following truth table, using propositional variables where previously we used specific propositions:

$p$	$q$	$p \ \& \ q$
T	T	T
T	F	F
F	T	F
F	F	F

There is no limit to the number of propositions we can conjoin to form a new proposition. “Roses are red and violets are blue; sugar is sweet and so are you” is a substitution instance of “ $p \ \& \ q \ \& \ r \ \& \ s$ .” We can also use

parentheses to group propositions. This last example could be treated as a substitution instance of " $(p \ \& \ q) \ \& \ (r \ \& \ s)$ "—that is, as a conjunction of two conjunctions. Later we will see that, just as in mathematics, parentheses can make an important difference to the meaning of a total proposition.

One cautionary note: The word "and" is not always used to connect two distinct sentences. Sometimes a sentence has to be rewritten for us to see that it is equivalent to a sentence of this form. For example,

Serena and Venus are tennis players.

is simply a short way of saying

Serena is a tennis player, and Venus is a tennis player.

At other times, the word "and" is *not* used to produce a conjunction of propositions. For example,

Serena and Venus are playing each other.

does *not* mean that

Serena is playing each other, and Venus is playing each other.

That does not even make sense, so the original sentence cannot express a conjunction of two propositions. Instead, it expresses a single proposition about two people taken as a group. Consequently, it should not be symbolized as " $p \ \& \ q$ ." Often, unfortunately, it is unclear whether a sentence expresses a conjunction of propositions or a single proposition about a group. The sentence

Serena and Venus are playing tennis.

could be taken either way. Maybe Serena and Venus are playing each other. If that is what it means, then the sentence expresses a single proposition about a group, so it should not be symbolized as " $p \ \& \ q$ ." But maybe Serena is playing one match, while Venus is playing another. If that would make it true, then the sentence expresses a conjunction of propositions, so it may be symbolized as " $p \ \& \ q$ ."

When a sentence containing the word "and" expresses the conjunction of two propositions, we will say that it expresses a *propositional conjunction*. When a sentence containing "and" does not express the conjunction of two propositions, we will say that it expresses a *nonpropositional conjunction*. In this chapter we are concerned only with sentences that express propositional conjunctions. A sentence should be translated into the symbolic form " $p \ \& \ q$ " only if it expresses a propositional conjunction. There is no mechanical procedure that can be followed to determine whether a certain sentence expresses a conjunction of two propositions. You must think carefully about what the sentence means and about the context in which that sentence is used. This takes practice.

**EXERCISE I**

The proposition “The night is young, and you’re so beautiful” is a substitution instance of which of the following propositional forms?

- |                 |                          |
|-----------------|--------------------------|
| 1. $p$          | 5. $p \ \& \ q \ \& \ r$ |
| 2. $q$          | 6. $p \ \& \ p$          |
| 3. $p \ \& \ q$ | 7. $p$ or $q$            |
| 4. $p \ \& \ r$ |                          |

**EXERCISE II**

Which of the following propositions is a substitution instance of “ $p \ \& \ q \ \& \ q'$ ”?

1. The night is young, and you’re so beautiful, and my flight leaves in thirty minutes.
2. The night is young, and you’re so beautiful, and my flight leaves in thirty minutes, and my flight leaves in thirty minutes.
3. You’re so beautiful, and you’re so beautiful, and you’re so beautiful.

**EXERCISE III**

For each of the following propositions, give three different propositional forms of which that proposition is a substitution instance.

1. The night is young, and you’re so beautiful, and my flight leaves in thirty minutes.
2. The night is young, and you’re so beautiful, and you’re so beautiful.

**EXERCISE IV**

Indicate whether each of the following sentences expresses a propositional conjunction or a nonpropositional conjunction—that is, whether or not it expresses a conjunction of two propositions. If the sentence could be either, then specify a context in which it would naturally be used to express a propositional conjunction and a different context in which it would naturally be used to express a nonpropositional conjunction.

1. A Catholic priest married John and Mary.
2. Fred had pie and ice cream for dessert.
3. The winning presidential candidate rarely loses both New York and California.

4. Susan got married and had a child.
5. Jane speaks both French and English.
6. Someone who speaks both French and English is bilingual.
7. Ken and Naomi are two of my best friends.
8. Miranda and Nick cooked dinner.
9. I doubt that John is poor and happy.

Now we can look at an argument involving conjunction. Here is one that is ridiculously simple:

Harry is short and John is tall.  
 $\therefore$  Harry is short.

This argument is obviously valid. But why is it valid? Why does the conclusion follow from the premise? The answer in this case seems obvious, but we will spell it out in detail as a guide for more difficult cases. Suppose we replace these particular propositions with propositional forms, using a different variable for each distinct proposition throughout the argument. This yields what we will call an *argument form*. For example:

$p \ \& \ q$   
 $\therefore p$

This is a pattern for endlessly many arguments, each of which is called a substitution instance of this argument form. Every argument that has this general form will also be valid. It really does not matter which propositions we put into this schema; the resulting argument will be valid—so long as we are careful to substitute the same proposition for the same variable throughout.

Let's pursue this matter further. If an argument has true premises and a false conclusion, then we know at once that it is invalid. But in saying that an argument is *valid*, we are not only saying that it does not have true premises and a false conclusion; we are also saying that the argument *cannot* have a false conclusion when the premises are true. Sometimes this is true because the argument has a structure or form that rules out the very possibility of true premises and a false conclusion. We can appeal to the notion of an argument form to make sense of this idea. A somewhat more complicated truth table will make this clear:

		PREMISE	CONCLUSION
$p$	$q$	$p \ \& \ q$	$p$
T	T	T	T
T	F	F	T
F	T	F	F
F	F	F	F

The first two columns give all the combinations for the truth values of the propositions that we might substitute for “ $p$ ” and “ $q$ .” The third column gives the truth value of the premise for each of these combinations. (This column is the same as the definition for “ $\&$ ” given above.) Finally, the fourth column gives the truth value for the conclusion for each combination. (Here, of course, this merely involves repeating the first column. Later on, things will become more complicated and interesting.) If we look at this truth table, we see that no matter how we make substitutions for the variables, we never have a case in which the premise is true and the conclusion is false. In the first line, the premise is true and the conclusion is also true. In the remaining three lines, the premise is not true, so the possibility of the premise being true and the conclusion false does not arise.

Here it is important to remember that a valid argument can have false premises, for one proposition can follow from another proposition that is false. Of course, an argument that is sound cannot have a false premise, because a sound argument is defined as a valid argument with true premises. But our subject here is validity, not soundness.

Let’s summarize this discussion. In the case we have examined, validity depends on the form of an argument and not on its particular content. A first principle, then, is this:

An *argument* is valid if it is an instance of a valid argument form.

Hence, the argument “Harry is short and John is tall; therefore, Harry is short” is valid because it is an instance of the valid argument form “ $p \ \& \ q; \therefore p$ .”

Next we must ask what makes an argument form valid. The answer to this is given in this principle:

An *argument form* is valid if and only if it has no substitution instances in which the premises are true and the conclusion is false.

We have just seen that the argument form “ $p \ \& \ q; \therefore p$ ” passes this test. The truth table analysis showed that. Incidentally, we can use the same truth table to show that the following argument is valid:

John is tall.	$p$
Harry is short.	$q$
$\therefore$ John is tall and Harry is short.	$\therefore p \ \& \ q$

The argument on the left is a substitution instance of the argument form on the right. A glance at the truth table will show that there can be no cases for which all the premises could be true and the conclusion false. This pretty well covers the logical properties of conjunction.

Notice that we have not said that *every* argument that is valid is so in virtue of its form. There may be arguments in which the conclusion follows from the premises but we cannot show how the argument’s validity is a

matter of logical form. There are, in fact, some obviously valid arguments that have yet to be shown to be valid in terms of their form. Explaining validity by means of logical form has long been an ideal of logical theory, but there are arguments—many of them quite common—where this ideal has yet to be adequately fulfilled. Many arguments in mathematics fall into this category. At present, however, we will only consider arguments in which the strategy we used for analyzing conjunction continues to work.

**EXERCISE V**

Are the following arguments valid by virtue of their propositional form? Why or why not?

1. Donald owns a tower in New York and a palace in Atlantic City.  
Therefore, Donald owns a palace in Atlantic City.
2. Tom owns a house. Therefore, Tom owns a house and a piece of land.
3. Ilsa is tall. Therefore, Ilsa is tall, and Ilsa is tall.
4. Bernie has a son and a daughter. Bernie has a father and a mother.  
Therefore, Bernie has a son and a mother.
5. Mary got married and had a child. Therefore, Mary had a child and got married.
6. Bess and Katie tied for MVP. Therefore, Bess tied for MVP.

**EXERCISE VI**

For each of the following claims, determine whether it is true or false. Defend your answers.

1. An argument that is a substitution instance of a valid argument form is always valid.
2. An argument that is a substitution instance of an invalid argument form is always invalid.
3. An invalid argument is always a substitution instance of an invalid argument form.

**DISCUSSION QUESTION**

Is a valid argument always a substitution instance of a valid argument form?  
Why or why not?

## DISJUNCTION

Just as we can form a conjunction of two propositions by using the connective “and,” we can form a *disjunction* of two propositions by using the connective “or,” as in the following compound sentence:

John will win or Harry will win.

Again, it is easy to see that the truth of this whole compound proposition depends on the truth of the component propositions. If they are both false, then the compound proposition is false. If just one of them is true, then the compound proposition is true. But suppose they are both true. What shall we say then?

Sometimes when we say “either-or,” we seem to rule out the possibility of both. When a waiter approaches your table and tells you, “Tonight’s dinner will be chicken or steak,” this suggests that you cannot have both. In other cases, however, it does not seem that the possibility of both is ruled out—for example, when we say to someone, “If you want to see tall mountains, go to California or Colorado.”

One way to deal with this problem is to say that the English word “or” has two meanings: one *exclusive*, which rules out both, and one *inclusive*, which does not rule out both. Another solution is to claim that the English word “or” always has the inclusive sense, but utterances with “or” sometimes conversationally imply the exclusion of both because of special features of certain contexts. It is, for example, our familiarity with common restaurant practices that leads us to infer that we cannot have both when the waiter says, “Tonight’s dinner will be chicken or steak.” If we may have both, then the waiter’s utterance would not be as informative as is required for the purpose of revealing our options, so it would violate Grice’s conversational rule of Quantity (as discussed in Chapter 2). That explains why the waiter’s utterance seems to exclude both.

Because such explanations are plausible, and because it is simpler as well as traditional to develop propositional logic with the inclusive sense of “or,” we will adopt that inclusive sense. Where necessary, we will define the exclusive sense using the inclusive sense as a starting point. Logicians symbolize disjunctions using the connective “ $\vee$ ” (called a *wedge*). The truth table for this connective has the following form:

$p$	$q$	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

We will look at some arguments involving this connective in a moment.

## NEGATION

With conjunction and disjunction, we begin with two propositions and construct a new proposition from them. There is another way in which we can

construct a new proposition from just one proposition—by *negating* it. Given the proposition “John is clever,” we can get a new proposition, “John is not clever,” simply by inserting the word “not” in the correct place in the sentence.

What, exactly, does the word “not” mean? This can be a difficult question to answer. Does it mean “nothing” or, maybe, “nothingness”? Although some respectable philosophers have sometimes spoken in this way, it is important to see that the word “not” does not stand for anything at all. It has an altogether different function in the language. To see this, think about how conjunction and disjunction work. Given two propositions, the word “and” allows us to construct another proposition that is true only when both original propositions are true, and false otherwise. With disjunction, given two propositions, the word “or” allows us to construct another proposition that is false only when both of the original propositions are false, and true otherwise. (Our truth table definitions reflect these facts.) Using these definitions as models, how should we define negation? A parallel answer is that the negation of a proposition is true just in the cases in which the original proposition is false, and it is false just in the cases in which original proposition is true. Using the symbol “ $\sim$ ” (called a *tilde*) to stand for negation, this gives us the following truth table definition:

$p$	$\sim p$
T	F
F	T

Negation might seem as simple as can be, but people quite often get confused by negations. If Diana says, “I could not breathe for a whole minute,” she might mean that there was a minute when something made her unable to breathe (maybe she was choking) or she might mean that she was able to hold her breath for a whole minute (say, to win a bet). If “ $A$ ” symbolizes “Diana could breathe sometime during this minute,” then “ $\sim A$ ” symbolizes the former claim (that Diana was unable to breathe for this minute). Consequently, the latter claim (that Diana could hold her breath for this minute) should not also be symbolized by “ $\sim A$ . ” Indeed, this interpretation of the original sentence is not a negation, even though the original sentence did include the word “not.” Moreover, some sentences are negations even though they do not include the word “not.” For example, “Nobody owns Mars” is the negation of “Somebody owns Mars.” If the latter is symbolized as “ $A$ ,” the former can be symbolized as “ $\sim A$ ,” even though the former does not include the word “not.”

The complexities of negation can be illustrated by noticing that the simple sentence “Everyone loves running” can include negation at four distinct places: “Not everyone loves running,” “Everyone does not love running,” “Everyone loves not running,” and the colloquial “Everyone loves running—not!” Some of these sentences can be symbolized in propositional logic as negations of “Everyone loves running,” but others cannot.

To determine whether a sentence can be symbolized as a negation in propositional logic, it is often useful to reformulate the sentence so that it starts

with “It is not the case that. . . .” For example, “I did none of the homework” would be reformulated as “It is not the case that I did any of the homework.” If the resulting sentence means the same as the original (as it does in this example), then the original sentence can be symbolized as a propositional negation. In contrast, “I promise not to leave you” means something very different from “It is not the case that I promise to leave you,” so “I promise not to leave you” should not be symbolized as a propositional negation.

Unfortunately, this test will not always work. There is no completely mechanical procedure for determining whether an English sentence can be symbolized as a negation. All you can do is think carefully about the sentence’s meaning and context. The best way to get good at this is to practice.

### EXERCISE VII

Explain the differences in meaning among “Not everyone loves running,” “Everyone does not love running,” “Everyone loves not running,” and “Everyone loves running—not!” For each, is it a negation of “Everyone loves running”? Why or why not?

### EXERCISE VIII

Negative terms or prefixes can often be interpreted in more than one way. Explain two ways to interpret each of the following sentences. Describe a context in which it would be natural to interpret it in each way.

1. You may not go to the meeting.
2. I cannot recommend him too highly.
3. He never thought he’d go to the Himalayas.
4. Have you not done all of your homework?
5. All of his friends are not students.
6. I will not go to some football games next season.
7. No smoking section available.
8. The lock on his locker was unlockable.

### EXERCISE IX

Put each of the following sentences in symbolic form. Be sure to specify exactly which sentence is represented by each capital letter, and pay special attention to the placement of the negation. If the sentence could be interpreted in more

than one way, symbolize each interpretation and describe a context in which it would be natural to interpret it in each way.

1. It won't rain tomorrow.
2. It might not rain tomorrow.
3. There is no chance that it will rain tomorrow.
4. I believe that it won't rain tomorrow.
5. Joe is not too smart or else he's very clever.
6. Kristin is not smart or rich.
7. Sometimes you feel like a nut; sometimes you don't. (from an advertisement for Mounds and Almond Joy candies, which are made by the same company and are exactly alike except that only one of them has a nut)

### PROCESS OF ELIMINATION

Using only negation and disjunction, we can analyze the form of one common pattern of reasoning, which is called *process of elimination* or, more technically, *disjunctive syllogism*. As an example, consider this argument:

She is sitting alone and talking, so she must be either talking on a phone I don't see or talking to herself. She is clearly not talking to herself, since she's not crazy. So she must be talking on her phone.

After trimming off assurances and subarguments that support the premises, the core of this argument can be put in standard form:

- (1) She is either talking to herself or talking on a phone.
  - (2) She is not talking to herself.
- ∴ (3) She is talking on a phone. (from 1–2)

This core argument is then an instance of this argument form:

$$\begin{array}{c} 1. p \vee q \\ \sim p \\ \hline \therefore q \end{array}$$

It does not matter if we change the order of the disjuncts so that the first premise is "She must be either talking on a phone or talking to herself." Then the argument takes this form:

$$\begin{array}{c} 2. p \vee q \\ \sim q \\ \hline \therefore p \end{array}$$

Both of these argument forms are valid, so the core of the original argument is also valid.

**EXERCISE X**

Explain why argument forms 1–2 are valid. Use common language that would be understandable to someone who has not read this chapter.

Process of elimination is sometimes confused with a similar but crucially different pattern of reasoning, which can be called *affirming a disjunct*. This pattern includes both of these forms:

$$\begin{array}{c} 3. \ p \vee q \\ \underline{p} \\ \therefore \sim q \end{array}$$

$$\begin{array}{c} 4. \ p \vee q \\ \underline{q} \\ \therefore \sim p \end{array}$$

These forms of argument are invalid. This can be shown by the following single instance:

She is either talking to herself or talking on a phone.

She is talking to herself.

$\therefore$  She is not talking on a phone.

This argument might seem valid if one assumes that she cannot talk on the phone while talking to herself. The premises, however, do not specify that she cannot do both at once. If she mumbles a few quick words to herself in the midst of talking on the phone, then the premises are both true and the conclusion is false.

**EXERCISE XI**

Give other instances of argument forms 3–4 that are not valid. Explain why these instances are invalid and why they show that the general argument form is invalid.

**HOW TRUTH-FUNCTIONAL CONNECTIVES WORK**

We have now defined conjunction, disjunction, and negation. That, all by itself, is sufficient to complete the branch of modern logic called propositional logic. The definitions themselves may seem peculiar. They do not look like the definitions we find in a dictionary. But the form of these definitions is important, for it tells us something interesting about the character of such words as “and,” “or,” and “not.” Two things are worth noting: (1) These expressions are used to construct a new proposition from old ones; (2) the newly constructed proposition is always a *truth function* of the original propositions—that is, the truth value of the new proposition is always

determined by the truth value of the original propositions. For this reason, these connectives are called *truth-functional connectives*. (Of course, with negation, we start with a *single* proposition, so there are not really two things to connect.) For example, suppose that “A” and “B” are two true propositions and “G” and “H” are two false propositions. We can then determine the truth values of more complex propositions built from them using conjunction, disjunction, and negation. Sometimes the correct assignment is obvious at a glance:

$A \ \& \ B$	True
$A \ \& \ G$	False
$\sim A$	False
$\sim G$	True
$A \vee H$	True
$G \vee H$	False
$\sim A \ \& \ G$	False

As noted earlier, parentheses can be used to distinguish groupings. Sometimes the placement of parentheses can make an important difference, as in the following two expressions:

$$\begin{aligned} &\sim A \ \& \ G \\ &\sim(A \ \& \ G) \end{aligned}$$

Notice that in the first expression the negation symbol applies only to the proposition “A,” whereas in the other expression it applies to the entire proposition “A & G.” Thus, the first expression above is false, and the second expression is true. Only the second expression can be translated as “Not both A and G.” Both of these expressions are different from “ $\sim A \ \& \ \sim G$ ,” which means “Neither A nor G.”

As expressions become more complex, we reach a point where it is no longer obvious how the truth values of the component propositions determine the truth value of the entire proposition. Here a regular procedure is helpful. The easiest method is to fill in the truth values of the basic propositions and then, step-by-step, make assignments progressively wider, going from the inside out. For example:

$$\begin{aligned} &\sim((A \vee G) \ \& \ \sim(\sim H \ \& \ B)) \\ &\sim((T \vee F) \ \& \ \sim(\sim F \ \& \ T)) \\ &\sim((T \vee F) \ \& \ \sim(T \ \& \ T)) \\ &\quad \sim(T \ \& \ \sim(T)) \\ &\quad \sim(T \ \& \ F) \\ &\quad \sim(F) \\ &\quad T \end{aligned}$$

With a little practice, you can master this technique in dealing with other very complex examples.

**EXERCISE XII**

Given that “*A*,” “*B*,” and “*C*” are true propositions and “*X*,” “*Y*,” and “*Z*” are false propositions, determine the truth values of the following compound propositions:

- |                               |  |
|-------------------------------|--|
| 1. $\sim X \vee Y$            | 9. $\sim(A \vee(Z \vee X))$  |
| 2. $\sim(X \vee Y)$           | 10. $\sim(A \vee \sim(Z \vee X))$                                  |
| 3. $\sim(Z \vee Z)$           | 11. $\sim A \vee \sim(Z \vee X)$                                   |
| 4. $\sim(Z \vee \sim Z)$      | 12. $\sim Z \vee(Z \& A)$  |
| 5. $\sim \sim(A \vee B)$      | 13. $\sim(Z \vee(Z \& A))$   |
| 6. $(A \vee Z) \& B$          | 14. $\sim((Z \vee Z) \& A)$  |
| 7. $(A \vee X) \& (B \vee Z)$ | 15. $A \vee ((\sim B \& C) \vee \sim(\sim B \vee \sim(Z \vee B)))$ |
| 8. $(A \& Z) \vee(B \& Z)$    | 16. $A \& ((\sim B \& C) \wedge \sim(\sim B \vee \sim(Z \vee B)))$ |

**TESTING FOR VALIDITY**

What is the point of all this? In everyday life, we rarely run into an expression as complicated as the one in our example at the end of the previous section. Our purpose here is to sharpen our sensitivity to how truth-functional connectives work and then to express our insights in clear ways. This is important because the validity of many arguments depends on the logical features of these truth-functional connectives. We can now turn directly to this subject.

Earlier we saw that every argument with the form “*p & q; ∴ p*” will be valid. This is obvious in itself, but we saw that this claim could be justified by an appeal to truth tables. A truth table analysis shows us that an argument with this form can never have an instance in which the premise is true and the conclusion is false. We can now apply this same technique to arguments that are more complex. In the beginning, we will examine arguments that are still easy to follow without the use of technical help. In the end, we will consider some arguments that most people cannot follow without guidance.

Consider the following argument:

Valerie is either a doctor or a lawyer.

Valerie is neither a doctor nor a stockbroker.

∴ Valerie is a lawyer.

We can use the following abbreviations:

*D* = Valerie is a doctor.

*L* = Valerie is a lawyer.

*S* = Valerie is a stockbroker.

Using these abbreviations, the argument and its counterpart argument form look like this:

$$\begin{array}{c} D \vee L \\ \sim(D \vee S) \\ \hline \therefore L \end{array} \qquad \begin{array}{c} p \vee q \\ \sim(p \vee r) \\ \hline \therefore q \end{array}$$

The expression on the right gives the argument form of the argument presented on the left. To test the argument for validity, we ask whether the argument form is valid. The procedure is cumbersome, but perfectly mechanical:

			PREMISE	PREMISE	CONCLUSION	
<i>p</i>	<i>q</i>	<i>r</i>	$(p \vee q)$	$(p \vee r)$	$\sim(p \vee r)$	<i>q</i>
T	T	T	T	T	F	T
T	T	F	T	T	F	T
T	F	T	T	T	F	F
T	F	F	T	T	F	F
F	T	T	T	T	F	T
F	T	F	T	F	T	T
F	F	T	F	T	F	F
F	F	F	F	F	T	F

Notice that there is only one combination of truth values for which both premises are true, and in that case the conclusion is true as well. So the original argument is valid because it is an instance of a valid argument form—that is, an argument form with no substitution instances for which true premises are combined with a false conclusion.

This last truth table may need some explaining. First, why do we get eight rows in this truth table where before we got only four? The answer to this is that we need to test the argument form for *every possible combination of truth values* for the component propositions. With two variables, there are four possible combinations: (TT), (TF), (FT), and (FF). With three variables, there are eight possible combinations: (TTT), (TTF), (TFT), (TFF), (FTT), (FTF), (FFT), and (FFF). The general rule is this: If an argument form has *n* variables, the truth table used in its analysis must have  $2^n$  rows. For four variables there will be sixteen rows; for five variables, thirty-two rows; for six variables, sixty-four rows; and so on. You can be sure that you capture all possible combinations of truth values by using the following pattern in constructing the columns of your truth table under each individual variable:

First column	Second column	Third column . . .
First half Ts, second half Fs.	First quarter Ts, second quarter Fs, and so on.	First eighth Ts, second eighth Fs, and so on.

A glance at the earlier examples in this chapter will show that we have been using this pattern, and it is the standard way of listing the possibilities.

Of course, as soon as an argument becomes at all complex, these truth tables become very large indeed. But there is no need to worry about this, because we will not consider arguments with many variables. Those who do so turn to a computer for help.

The style of the truth table above is also significant. The premises are plainly labeled, and so is the conclusion. A line is drawn under every row in which the premises are all true. (In this case, there is only one such row—row 6.) If the conclusion on this line is also true, it is marked “OK.” If every line in which the premises are all true is OK, then the argument form is valid. Marking all this may seem rather childish, but it is worth doing. First, it helps guard against mistakes. More importantly, it draws one’s attention to the purpose of the procedure being used. Cranking out truth tables without understanding what they are about—or even why they might be helpful—does not enlighten the mind or elevate the spirit.

For the sake of contrast, we can next consider an invalid argument:

- (1) Valerie is either a doctor or a lawyer.
  - (2) Valerie is not both a lawyer and a stockbroker.
- $\therefore$  (3) Therefore, Valerie is a doctor.

Using the same abbreviations as earlier, this becomes:

$$\begin{array}{c} D \vee L \\ \sim(L \& S) \\ \hline \therefore D \end{array} \qquad \begin{array}{c} p \vee q \\ \sim(q \& r) \\ \hline \therefore p \end{array}$$

The truth table for this argument form looks like this:

			PREMISE	PREMISE	CONCLUSION	
p	q	r	$(p \vee q)$	$(q \& r)$	$\sim(q \& r)$	p
T	T	T	T	T	F	T
T	T	F	T	F	T	T
T	F	T	T	F	T	T
T	F	F	T	F	T	T
F	T	T	T	T	F	F
F	T	F	T	F	T	F
F	F	T	F	F	T	F
F	F	F	F	F	T	F

This time, we find four rows in which all the premises are true. In three cases the conclusion is true as well, but in one of these cases (row 6), the conclusion is false. This line is marked “Invalid.” Notice that every line in which all of the premises are true is marked either as “OK” or as “Invalid.” If even one row is marked “Invalid,” then the argument form as a whole is invalid. The argument form under consideration is thus invalid, because it is possible for it to have a substitution instance in which all the premises are true and the conclusion is false.

The labeling not only shows *that* the argument form is invalid, it also shows *why* it is invalid. Each line that is marked “Invalid” shows a combination of truth values that makes the premises true and the conclusion false. Row 6 presents the combination in which Valerie is not a doctor, is a lawyer, and is not a stockbroker. With these assignments, it will be true that she is either a doctor or a lawyer (premise 1), and also true that she is not both a lawyer and a stockbroker (premise 2), yet false that she is a doctor (the conclusion). It is this possibility that shows why the argument form is not valid.

In sum, we can test a propositional argument form for validity by following these simple steps:

1. Provide a column for each premise and the conclusion.
2. Fill in truth values in each column.
3. Underline each row where all of the premises are true.
4. Mark each row “OK” if the conclusion is true on that row.
5. Mark each row “Invalid” if the conclusion is false on that row.
6. If any row is marked “Invalid,” the argument form is invalid.
7. If no row is marked “Invalid,” the argument form is valid.

#### EXERCISE XIII

Using the truth table technique outlined above, show that argument forms 1–2 in the above section on process of elimination are valid and that argument forms 3–4 in the same section are invalid.

#### EXERCISE XIV

Is the following argument valid in our technical sense? Explain why or why not. Could it be sound? Explain why or why not.

- (1) Frogs are green.
- (2) Frogs are not green.
- ∴ (3) I am president. (from 1–2)

#### EXERCISE XV

Using the truth table technique outlined above, test the following argument forms for validity:

$$\begin{array}{ll}
 \begin{array}{l}
 1. \sim p \vee q \\
 \quad \underline{p} \\
 \therefore \sim q
 \end{array} &
 \begin{array}{l}
 2. \sim(p \vee q) \\
 \quad \underline{\quad} \\
 \therefore \sim q
 \end{array}
 \end{array}$$

(continued)

3. $\sim(p \vee q)$	7. $(p \& q) \vee (p \& r)$
$\frac{p}{\therefore q}$	$\therefore \overline{p \& (q \vee r)}$
4. $\sim(p \vee q)$	8. $\underline{(p \vee q) \& (p \vee r)}$
$\frac{p}{\therefore r}$	$\therefore p \& (q \vee r)$
5. $\sim(p \& q)$	9. $\underline{p \& q}$
$\frac{q}{\therefore \sim p}$	$\therefore (p \vee r) \& (q \vee r)$
6. $\sim(p \& q)$	10. $\underline{p \vee q}$
$\frac{\sim q}{\therefore p}$	$\therefore (p \& r) \vee (q \& r)$

## SOME FURTHER CONNECTIVES

We have developed the logic of propositions using only three basic notions corresponding (perhaps roughly) to the English words “and,” “or,” and “not.” Now let us go back to the question of the two possible senses of the word “or”: one exclusive and the other inclusive. Sometimes “or” seems to rule out the possibility that both alternatives are true; at other times “or” seems to allow this possibility. This is the difference between exclusive and inclusive disjunction.

Suppose we use the symbol “ $\underline{\vee}$ ” to stand for exclusive disjunction. This is the same as the symbol for inclusive disjunction except that it is underlined. (After this discussion, we will not use it again.) We could then give two truth table definitions, one for each of these symbols:

		INCLUSIVE	EXCLUSIVE
$p$	$q$	$p \vee q$	$p \underline{\vee} q$
T	T	T	F
T	F	T	T
F	T	T	T
F	F	F	F

We could also define this new connective in the following way:

$$(p \underline{\vee} q) = (\text{by definition}) ((p \vee q) \& \sim(p \& q))$$

It is not hard to see that the expression on the right side of this definition captures the force of exclusive disjunction. Because we can always define exclusive disjunction when we want it, there is no need to introduce it into our system of basic notions.

**EXERCISE XVI**

Construct a truth table analysis of the expression on the right side of the preceding definition, and compare it with the truth table definition of exclusive disjunction.

**EXERCISE XVII**

Use truth tables to test the following argument forms for validity:

- |                             |                         |
|-----------------------------|-------------------------|
| 1. $p$ _____                | 4. $\neg(p \& q)$ _____ |
| $\therefore p \leq q$       | $\therefore p \leq q$   |
| 2. $p \vee q$ _____         | 5. $p \leq q$ _____     |
| $p$ _____                   | $\therefore p \vee q$   |
| $\therefore \neg q$         | 6. $p \vee q$ _____     |
| 3. $p \& q$ _____           | $\therefore p \leq q$   |
| $\therefore \neg(p \leq q)$ |                         |

Actually, in analyzing arguments we have been defining new logical connectives without thinking about it much. For example, “not both  $p$  and  $q$ ” was symbolized as “ $\neg(p \& q)$ .” “Neither  $p$  nor  $q$ ” was symbolized as “ $\neg(p \vee q)$ .” Let us look more closely at the example “ $\neg(p \vee q)$ .” Perhaps we should have symbolized it as “ $\neg p \& \neg q$ .” In fact, we could have used this symbolization, because the two expressions amount to the same thing. Again, this may be obvious, but we can prove it by using a truth table in yet another way. Compare the truth table analysis of these two expressions:

$p$	$q$	$\neg p$	$\neg q$	$\neg p \& \neg q$	$(p \vee q)$	$\neg(p \vee q)$
T	T	F	F	F	T	F
T	F	F	T	F	T	F
F	T	T	F	F	T	F
F	F	T	T	T	F	T

Under “ $\neg p \& \neg q$ ” we find the column (FFFT), and we find the same sequence under “ $\neg(p \vee q)$ .” This shows that, for every possible substitution we make, these two expressions will yield propositions with the same truth value. We will say that these propositional forms are *truth-functionally equivalent*. The above table also shows that the expressions “ $\neg q$ ” and “ $\neg p \& \neg q$ ” are *not* truth-functionally equivalent, because the columns underneath these two expressions differ in the second row, so some substitutions into these expressions will not yield propositions with the same truth value.

Given the notion of truth-functional equivalence, the problem of more than one translation can often be solved. If two translations of a sentence are

truth-functionally equivalent, then it does not matter which one we use in testing for validity. Of course, some translations will seem more natural than others. For example, “ $p \vee q$ ” is truth-functionally equivalent to

$$\sim((\sim p \& \sim q) \& (\sim q \vee \sim p))$$

Despite this equivalence, the first form of expression is obviously more natural than the second when translating sentences, such as “It is either cloudy or sunny.”

### EXERCISE XVIII

Use truth tables to test which of the following propositional forms are truth-functionally equivalent to each other:

1.  $\sim(p \vee q)$
2.  $\sim(\sim p \vee \sim q)$
3.  $\sim p \& \sim q$
4.  $p \& q$

### EXERCISE XIX

Use truth tables to determine whether the expressions in each of the following pairs are truth-functionally equivalent:

- |  |   |
|--|---|
| 1. “ $p$ ” and “ $p \& p$ ”                        | 9. “ $\sim(p \vee q)$ ” and “ $\sim p \vee q$ ”                       |
| 2. “ $p$ ” and “ $p \vee p$ ”                      | 10. “ $\sim(p \vee q)$ ” and “ $\sim p \& \sim q$ ”                   |
| 3. “ $p \vee \sim p$ ” and “ $\sim(p \& \sim p)$ ” | 11. “ $\sim\sim(p \vee q)$ ” and “ $\sim\sim p \& \sim\sim q$ ”       |
| 4. “ $p$ ” and “ $p \& (q \vee \sim q)$ ”          | 12. “ $\sim(p \& q)$ ” and “ $\sim p \vee q$ ”                        |
| 5. “ $p$ ” and “ $p \& (q \& \sim q)$ ”            | 13. “ $\sim\sim(p \& q)$ ” and “ $\sim\sim p \vee \sim\sim q$ ”       |
| 6. “ $p$ ” and “ $p \vee (q \& \sim q)$ ”          | 14. “ $\sim\sim p \& \sim\sim q$ ” and “ $\sim(\sim p \& \sim q)$ ”   |
| 7. “ $p \& (q \vee r)$ ” and “ $p \vee (q \& r)$ ” | 15. “ $\sim\sim p \& \sim\sim q$ ” and “ $\sim(\sim p \vee \sim q)$ ” |
| 8. “ $p \& (q \& r)$ ” and “ $(p \& q) \& r$ ”     | 16. “ $p \& \sim\sim q$ ” and “ $\sim\sim p \& q$ ”                   |

## CONDITIONALS

So far in this chapter we have seen that by using conjunction, disjunction, and negation, it is possible to construct compound propositions out of simple propositions. A distinctive feature of compound propositions constructed in these three ways is that the truth of the compound proposition is always a function of the truth of its component propositions. Thus, these three notions allow us to construct truth-functionally compound propositions. Some arguments depend for their validity simply on these truth-functional

connectives. When this is so, it is possible to test for validity in a purely mechanical way. This can be done through the use of truth tables. Thus, in this area at least, we are able to give a clear account of validity and to specify exact procedures for testing for validity.

This truth-functional approach might seem problematic in another area: *conditionals*. We will argue that an important group of conditionals can be handled in much the same way as negation, conjunction, and disjunction. We separate conditionals from the other connectives only because a truth-functional treatment of conditionals is more controversial and faces problems that are instructive.

Conditionals have the form “If \_\_\_\_\_, then \_\_\_\_\_. ” What goes in the first blank of this pattern is called the *antecedent* of the conditional; what goes in the second blank is called its *consequent*. Sometimes conditionals appear in the indicative mood:

If it rains, then the crop will be saved.

Sometimes they occur in the subjunctive mood:

If it had rained, then the crop would have been saved.

There are also conditional imperatives:

If a fire breaks out, then call the fire department first!

And there are conditional promises:

If you get into trouble, then I promise to help you.

Indeed, conditionals get a great deal of use in our language, often in arguments. It is important, therefore, to understand them.

Unfortunately, there is no general agreement among experts concerning the correct way to analyze conditionals. We will simplify matters and avoid some of these controversies by considering only indicative conditionals. We will not examine conditional imperatives, conditional promises, or subjunctive conditionals. Furthermore, at the start, we will examine only what we will call *propositional conditionals*. We get a propositional conditional by substituting indicative sentences that express propositions—something either true or false—into the schema “If \_\_\_\_\_, then \_\_\_\_\_. ” Or, to use technical language already introduced, a propositional conditional is a substitution instance of “If *p*, then *q*” in which “*p*” and “*q*” are propositional variables. Of the four conditional sentences listed above, only the first is clearly a propositional conditional.

Even if we restrict our attention to propositional conditionals, this will not avoid all controversy. Several competing theories claim to provide the correct analysis of propositional conditionals, and no consensus has been reached concerning which is right. It may seem surprising that theorists disagree about such a simple and fundamental notion as the if-then construction, but they do. In what follows, we will first describe the most standard treatment of propositional conditionals, and then consider alternatives to it.

## TRUTH TABLES FOR CONDITIONALS

For conjunction, disjunction, and negation, the truth table method provides an approach that is at once plausible and effective. A propositional conditional is also compounded from two simpler propositions, and this suggests that we might be able to offer a truth table definition for these conditionals as well. What should the truth table look like? When we try to answer this question, we get stuck almost at once, for it is unclear how we should fill in the table in three out of four cases.

$p$	$q$	If $p$ , then $q$
T	T	?
T	F	F
F	T	?
F	F	?

It seems obvious that a conditional cannot be true if the antecedent is true and the consequent is false. We record this by putting “F” in the second row. But suppose “ $p$ ” and “ $q$ ” are replaced by two arbitrary true propositions—say, “Two plus two equals four” and “Chile is in South America.” Consider what we shall say about the conditional:

If two plus two equals four, then Chile is in South America.

This is a *very* strange statement, because the arithmetical remark in the antecedent does not seem to have anything to do with the geographical remark in the consequent. So this conditional is odd—indeed, extremely odd—but is it true or false? At this point, a reasonable response is bafflement.

Consider the following argument, which is intended to solve all these problems by providing reasons for assigning truth values in each row of the truth table. First, it seems obvious that, if “If  $p$ , then  $q$ ” is true, then it is not the case that both “ $p$ ” is true and “ $q$ ” is false. That in turn means that “ $\sim(p \ \& \ \sim q)$ ” must be true. The following, then, seems to be a valid argument form:

$$\begin{array}{c} \text{If } p, \text{ then } q. \\ \hline \therefore \sim(p \ \& \ \sim q) \end{array}$$

Second, we can also reason in the opposite direction. Suppose we know that “ $\sim(p \ \& \ \sim q)$ ” is true. For this to be true, “ $p \ \& \ \sim q$ ” must be false. We know this from the truth table definition of negation. Next let us suppose that “ $p$ ” is true. Then “ $\sim q$ ” must be false. We know this from the truth table definition of conjunction. Finally, if “ $\sim q$ ” is false, then “ $q$ ” itself must be true. This line of reasoning is supposed to show that the following argument form is valid:

$$\begin{array}{c} \sim(p \ \& \ \sim q) \\ \hline \therefore \text{If } p, \text{ then } q. \end{array}$$

The first step in the argument was intended to show that we can validly derive “ $\sim(p \ \& \ \sim q)$ ” from “If  $p$ , then  $q$ . ” The second step was intended to show that the

derivation can be run in the other direction as well. But if each of these expressions is derivable from the other, this suggests that they are equivalent. We use this background argument as a justification for the following definition:

If  $p$ , then  $q$  = (by definition) not both  $p$  and not  $q$ .

We can put this into symbolic notation using “ $\supset$ ” (called a *horseshoe*) to symbolize the conditional connective:

$p \supset q$  = (by definition)  $\sim(p \ \& \ \sim q)$

Given this definition, we can now construct the truth table for propositional conditionals. It is simply the truth table for “ $\sim(p \ \& \ \sim q)$ ”:

$p$	$q$	$\sim(p \ \& \ \sim q)$	$p \supset q$	$\sim p \vee q$
T	T	T	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

Notice that “ $\sim(p \ \& \ \sim q)$ ” is also truth-functionally equivalent to the expression “ $\sim p \vee q$ .” We have cited it here because “ $\sim p \vee q$ ” has traditionally been used to define “ $p \supset q$ .” For reasons that are now obscure, when a conditional is defined in this truth-functional way, it is called a *material conditional*.

Let’s suppose, for the moment, that the notion of a material conditional corresponds exactly to our idea of a propositional conditional. What would follow from this? The answer is that we could treat conditionals in the same way in which we have treated conjunction, disjunction, and negation. A propositional conditional would be just one more kind of truth-functionally compound proposition capable of definition by truth tables. Furthermore, the validity of arguments that depend on this notion (together with conjunction, disjunction, and negation) could be settled by appeal to truth table techniques. Let us pause for a moment to examine this.

One of the most common patterns of reasoning is called *modus ponens*. It looks like this:

$$\begin{array}{c} \text{If } p, \text{ then } q. \\ p \\ \hline \therefore q \end{array} \qquad \begin{array}{c} p \supset q \\ p \\ \hline \therefore q \end{array}$$

The truth table definition of a material conditional shows at once that this pattern of argument is valid:

PREMISE		PREMISE		CONCLUSION
$p$	$q$	$p \supset q$	$q$	
T	T	T	T	OK
T	F	F	F	
F	T	T	T	
F	F	T	F	

**EXERCISE XX**

The argument form called *modus tollens* looks like this:

$$\begin{array}{c} p \supset q \\ \sim q \\ \hline \therefore \sim p \end{array}$$

Use truth tables to show that this argument form is valid.

**Farcus**

by David Waisglass  
Gordon Coulthart



Farcus/LaughingStock Licensing Inc.

© 1997 Farcus Cartoons

**"So, I say if it's not worth doing well,  
it's not worth doing at all."**

These same techniques allow us to show that one of the traditional fallacies is, indeed, a fallacy. It is called the fallacy of *denying the antecedent*, and it has this form:

$$\begin{array}{c} p \supset q \\ \sim p \\ \hline \therefore \sim q \end{array}$$

The truth table showing the invalidity of this argument form looks like this:

		PREMISE	PREMISE	CONCLUSION	
$p$	$q$	$p \supset q$	$\sim p$	$\sim q$	
T	T	T	F	F	
T	F	F	F	T	
F	T	T	T	F	Invalid
F	F	T	T	T	OK

#### EXERCISE XXI

A second standard fallacy is called *affirming the consequent*. It looks like this:

$$\begin{array}{c} p \supset q \\ q \\ \hline \therefore p \end{array}$$

Use truth tables to show that this argument form is invalid.

#### EXERCISE XXII

In his radio address to the nation on April 17, 1982, President Ronald Reagan argued that the United States should not accept a treaty with the Soviet Union that would mutually freeze nuclear weapons at current levels, because he believed that the United States had fallen behind. Here is a central part of his argument:

It would be wonderful if we could restore the balance of power with the Soviet Union without increasing our military power. And, ideally, it would be a long step towards assuring peace if we could have significant and verifiable reductions of arms on both sides. But let's not fool ourselves. The Soviet Union will not come to any conference table bearing gifts. Soviet negotiators will not make unilateral concessions. To achieve parity, we must make it plain that we have the will to achieve parity by our own effort.

Put Reagan's central argument into standard form. Then symbolize it and its form. Does his argument commit any fallacy? If so, identify it.

The relations among these last four argument forms can be seen in this diagram:

	Antecedent	Consequent
<b>Affirming</b>	Affirming the Antecedent = <i>Modus Ponens</i> (valid)	Affirming the Consequent (invalid)
<b>Denying</b>	Denying the Antecedent (invalid)	Denying the Consequent = <i>Modus Tollens</i> (valid)

Another argument form that has been historically significant is called a *hypothetical syllogism*:

$$\begin{array}{c} p \supset q \\ q \supset r \\ \hline \therefore p \supset r \end{array}$$

Because we are dealing with an argument form containing three variables, we must perform the boring task of constructing a truth table with eight rows:

			PREMISE	PREMISE	CONCLUSION	
<i>p</i>	<i>q</i>	<i>r</i>	<i>p</i> $\supset$ <i>q</i>	<i>q</i> $\supset$ <i>r</i>	<i>p</i> $\supset$ <i>r</i>	
T	T	T	T	T	T	OK
T	T	F	T	F	F	
T	F	T	F	T	T	
T	F	F	F	T	F	
F	T	T	T	T	T	OK
F	T	F	T	F	T	
F	F	T	T	T	T	OK
F	F	F	T	T	T	OK

This is fit work for a computer, not for a human being, but it is important to see that it actually works.

Why is it important to see that these techniques work? Most people, after all, could see that hypothetical syllogisms are valid without going through all of this tedious business. We seem only to be piling boredom on top of triviality. This protest deserves an answer. Suppose we ask someone *why* he or she thinks that the conclusion follows from the premises in a hypothetical syllogism. The person might answer that anyone can see that—which, by the way, is false. Beyond this, he or she might say that it all depends on the meanings of the words or that it is all a matter of definition. But if we go on to ask, “which words?” and “what definitions?” then most people will fall silent. We have discovered that the validity of some arguments depends on the meanings of such words as “and,” “or,” “not,” and “if-then.” We have then gone on to give explicit definitions of these terms—definitions, by the way, that help us see how these terms function in an argument. Finally, by getting all these *simple* things right, we have produced what is called a *decision procedure* for determining the validity of every argument depending only on conjunctions, disjunctions, negations, and propositional conditionals. Our truth table techniques give us a mechanical procedure for settling questions of validity in this area. In fact, truth table techniques have practical applications, for example, in computer programming. But the important point here is that, through an understanding of how these techniques work, we can gain a deeper insight into the notion of validity.

**EXERCISE XXIII**

Two more classic, common, and useful argument forms combine conditionals with disjunction. Using truth tables, test them for validity.

*Constructive Dilemma*

$$\begin{array}{c} p \vee q \\ p \supset r \\ q \supset r \\ \hline \therefore r \end{array}$$

*Destructive Dilemma*

$$\begin{array}{c} \sim p \vee \sim q \\ r \supset p \\ r \supset q \\ \hline \therefore \sim r \end{array}$$

**EXERCISE XXIV**

Using the truth table techniques employed above, test the following argument forms for validity. (For your own entertainment, guess whether the argument form is valid or invalid before working it out.)

$$\begin{array}{c} 1. \ p \supset q \\ \therefore q \supset p \end{array}$$

$$\begin{array}{c} 8. \ p \supset (q \vee r) \\ \sim q \\ \sim r \\ \hline \therefore \sim p \end{array}$$

$$\begin{array}{c} 2. \ p \supset q \\ \therefore \sim q \supset \sim p \end{array}$$

$$\begin{array}{c} 9. \ (p \vee q) \supset r \\ \therefore p \supset r \end{array}$$

$$\begin{array}{c} 3. \ \sim q \supset \sim p \\ \therefore p \supset q \end{array}$$

$$\begin{array}{c} 10. \ (p \ \& \ q) \supset r \\ \therefore p \supset r \end{array}$$

$$\begin{array}{c} 4. \ p \supset q \\ q \supset r \\ \hline \therefore p \supset (q \ \& \ r) \end{array}$$

$$\begin{array}{c} 11. \ p \supset (q \supset r) \\ \therefore (p \ \& \ q) \supset r \end{array}$$

$$\begin{array}{c} 5. \ p \supset q \\ q \supset r \\ \sim r \\ \hline \therefore \sim p \end{array}$$

$$\begin{array}{c} 12. \ (p \ \& \ q) \supset r \\ \therefore p \supset (q \supset r) \end{array}$$

$$\begin{array}{c} 6. \ p \supset q \\ q \supset r \\ \hline \therefore \sim r \supset \sim p \end{array}$$

$$\begin{array}{c} 13. \ p \supset (q \supset r) \\ q \\ \sim r \\ \hline \therefore \sim p \end{array}$$

$$\begin{array}{c} 7. \ p \vee q \\ p \supset q \\ q \supset r \\ \hline \therefore r \end{array}$$

$$\begin{array}{c} 14. \ p \supset (q \supset r) \\ p \supset q \\ \hline \therefore r \end{array}$$

(continued)

15. 
$$\frac{(p \vee q) \ \& \ (p \vee r)}{\therefore \sim q}$$
16. 
$$\frac{(p \supset q) \ \& \ (p \supset \sim r)}{\therefore \sim p}$$
17. 
$$\frac{(p \vee q) \supset p}{\therefore \sim q}$$
18. 
$$\frac{(p \vee q) \supset (p \ \& \ q)}{\therefore (p \supset q) \ \& \ (q \supset p)}$$
19. 
$$\frac{(p \ \& \ q) \supset (p \ \& \ q)}{\therefore (p \supset q) \ \& \ (q \supset p)}$$
20. 
$$\frac{r}{\therefore (p \supset q) \vee (q \supset p)}$$

## LOGICAL LANGUAGE AND EVERYDAY LANGUAGE

Early in this chapter we started out by talking about such common words as “and” and “or,” and then we slipped over to talking about *conjunction* and *disjunction*. The transition was a bit sneaky, but intentional. To understand what is going on here, we can ask how closely these logical notions we have defined match their everyday counterparts. We will start with conjunction, and then come back to the more difficult question of conditionals.

At first sight, the match between conjunction as we have defined it and the everyday use of the word “and” may seem fairly bad. To begin with, in everyday discourse, we do not go about conjoining random bits of information. We do not say, for example, “Two plus two equals four and Chile is in South America.” We already know why we do not say such things, for unless the context is quite extraordinary, this is bound to violate the conversational rule of Relevance. But if we are interested in validity, the rule of Relevance—like all other conversational (or pragmatic) rules—is simply beside the point. When dealing with validity, we are interested in only one question: If the premises of an argument are true, must the conclusion be true as well? Conversational rules, as we saw in Chapter 2, do not affect truth.

The truth-functional notion of conjunction is also insensitive to another important feature of our everyday discourse: By reducing all conjunctions to their bare truth-functional content, the truth-functional notion often misses the argumentative point of a conjunction. As we saw in Chapter 3, each of the following remarks has a different force in the context of an argument:

The ring is beautiful, but expensive.

The ring is expensive, but beautiful.

These two remarks point in opposite directions in the context of an actual argument, but from a purely truth-functional point of view, we treat them as equivalent. We translate the first sentence as “*B & E*” and the second as

"E & B." Their truth-functional equivalence is too obvious to need proof. Similar oddities arise for all discounting terms, such as "although," "whereas," and "however."

It might seem that if formal analysis cannot distinguish an "and" from a "but," then it can hardly be of any use at all. This is not true. A formal analysis of an argument will tell us just one thing: whether the argument is valid or not. If we expect the analysis to tell us more than this, we will be sorely disappointed. It is important to remember two things: (1) We expect deductive arguments to be valid, and (2) usually we expect much more than this from an argument. To elaborate on the second point, we usually expect an argument to be sound as well as valid; we expect the premises to be true. Beyond this, we expect the argument to be informative, intelligible, convincing, and so forth. Validity, then, is an important aspect of an argument, and formal analysis helps us evaluate it. But validity is not the only aspect of an argument that concerns us. In many contexts, it is not even our chief concern.

We can now look at our analysis of conditionals, for here we find some striking differences between the logician's analysis and everyday use. The following argument forms are both valid:

$$\begin{array}{ll} 1. p & 2. \sim p \\ \therefore q \supset p & \therefore p \supset q \end{array}$$

#### EXERCISE XXV

Check the validity of the argument forms above using truth tables.

Though valid, both argument forms seem odd—so odd that they have actually been called *paradoxical*. The first argument form seems to say this: If a proposition is true, then it is *implied* by any proposition whatsoever. Here is an example of an argument that satisfies this argument form and is therefore valid:

Lincoln was president.

$\therefore$  If the moon is made of cheese, Lincoln was president.

This is a peculiar argument to call valid. First, we want to know what the moon has to do with Lincoln's having been president. Beyond this, how can his having been president depend on a blatant falsehood? We can give these questions even more force by noticing that even the following argument is valid:

Lincoln was president.

$\therefore$  If Lincoln was not president, then Lincoln was president.

Both arguments are instances of the valid argument form " $p; \therefore q \supset p$ ."

The other argument form is also paradoxical. It seems to say that a false proposition implies any proposition whatsoever. The following is an instance of this argument form:

Columbus was not president.

$\therefore$  If Columbus was president, then the moon is made of cheese.

Here it is hard to see what the falsehood that Columbus was president has to do with the composition of the moon.

At this point, nonphilosophers become impatient, whereas philosophers become worried. We started out with principles that seemed to be both obvious and simple. Now, quite suddenly, we are being overwhelmed with a whole series of peculiar results. What in the world has happened, and what should be done about it? Philosophers remain divided in the answers they give to these questions. The responses fall into two main categories: (1) Simply give up the idea that conditionals can be defined by truth-functional techniques and search for a different and better analysis of conditionals that avoids the difficulties involved in truth-functional analysis; or (2) take the difficult line and argue that there is nothing wrong with calling the aforementioned argument forms valid.

The first approach is highly technical and cannot be pursued in detail in this book, but the general idea is this: Instead of identifying "If  $p$ , then  $q$ " with "Not both  $p$  and not  $q$ ," identify it with "Not possibly both  $p$  and not  $q$ ." This provides a stronger notion of a conditional and avoids some—though not all—of the problems concerning conditionals. This theory is given a systematic development by offering a logical analysis of the notion of possibility. This branch of logic is called *modal* logic, and it has shown remarkable development in recent decades.

The second line has been taken by Paul Grice, whose theories played a prominent part in Chapter 2. He acknowledges—as anyone must—that the two argument forms above are decidedly odd. He denies, however, that this oddness has anything to do with *validity*. Validity concerns one thing and one thing only: a relationship between premises and conclusion. An argument is valid if the premises cannot be true without the conclusion being true as well. The above arguments are valid by this definition of "validity."

Of course, arguments can be defective in all sorts of other ways. Look at the first argument form: (1)  $p; \therefore q \supset p$ . Because " $q$ " can be replaced by any proposition (true or false), the rule of Relevance will often be violated. It is worth pointing out violations of the rule of Relevance, but, according to Grice, this issue has nothing to do with validity. Beyond this, arguments having this form can also involve violations of the rule of Quantity. A conditional will be true whenever the consequent is true. Given this, it does not matter to the truth of the whole conditional whether the antecedent is true

or false. Yet it can be misleading to use a conditional on the basis of this logical feature. For example, it would be misleading for a museum guard to say, "If you give me five dollars, then I will let you into the exhibition," when, in fact, he will admit you in any case. For Grice, this is misleading because it violates the rule of Quantity. Yet strictly speaking, it is not false. Strictly speaking, it is true.

The Grice line is attractive because, among other things, it allows us to accept the truth-functional account of conditionals, with all its simplicity. Yet sometimes it is difficult to swallow. Consider the following remark:

If God exists, then there is evil in the world.

If Grice's analysis is correct, even the most pious person will have to admit that this conditional is true provided only that he or she is willing to admit that there is evil in the world. Yet, this conditional plainly suggests that there is some connection between God's existence and the evil in the world—presumably, that is the point of connecting them in a conditional. The pious will wish to deny this suggestion. All the same, this connection is something that is conversationally implied, not asserted. So, once more, this conditional could be misleading—and therefore is in need of criticism and correction—but it is still, strictly speaking, true.

Philosophers and logicians have had various responses to Grice's position. No consensus has emerged on this issue. The authors of this book find it adequate, at least in most normal cases, and therefore have adopted it. This has two advantages: (1) The appeal to conversational rules fits in well with our previous discussions, and (2) it provides a way of keeping the logic simple and within the range of a beginning student. Other philosophers and logicians continue to work toward a definition superior to the truth table definition for indicative conditionals.

## OTHER CONDITIONALS IN ORDINARY LANGUAGE

So far we have considered only one form in which propositional conditionals appear in everyday language: the conditional "If *p*, then *q*." But propositional conditionals come in a variety of forms, and some of them demand careful treatment.

We can first consider the contrast between constructions using "if" and those using "only if":

1. I'll clean the barn if Hazel will help me.
2. I'll clean the barn only if Hazel will help me.

Adopting the following abbreviations:

*B* = I'll clean the barn

*H* = Hazel will help me

the first sentence is symbolized as follows:

$$H \supset B$$

Notice that in the prose version of item 1, the antecedent and consequent appear in reverse order; “*q if p*” means the same thing as “If *p*, then *q*.”

How shall we translate the second sentence? Here we should move slowly and first notice what seems uncontested: If Hazel does not help me, then I will not clean the barn. This is translated in the following way:

$$\sim H \supset \sim B$$

And that is equivalent to:

$$B \supset H$$

If this equivalence is not obvious, it can quickly be established using a truth table.

A more difficult question arises when we ask whether an implication runs the other way. When I say that I will clean the barn only if Hazel will help me, am I committing myself to cleaning the barn if she does help me? There is a strong temptation to answer the question “yes” and then give a fuller translation of item 2 in the following way:

$$(B \supset H) \ \& \ (H \supset B)$$

Logicians call such two-way implications *biconditionals*, and we will discuss them in a moment. But adding this second conjunct is almost surely a mistake, for we can think of parallel cases where we would not be tempted to include it. A government regulation might read as follows:

A student may receive a New York State Scholarship only if the student attends a New York State school.

From this it does not follow that anyone who attends a New York State school may receive a New York State Scholarship. There may be other requirements as well—for example, being a New York State resident.

Why were we tempted to use a biconditional in translating sentences containing the connective “only if”? Why, that is, are we tempted to think that the statement “I’ll clean the barn only if Hazel will help me” implies “If Hazel helps me, then I will clean the barn”? The answer turns on the notion of conversational implication first discussed in Chapter 2. If I am *not* going to clean the barn whether Hazel helps me or not, then it will be misleading—a violation of the rule of Quantity—to say that I will clean the barn only if Hazel helps me. For this reason, in many contexts, the *use* of a sentence of the form “*p* only if *q*” will conversationally imply a commitment to “*p* if and only if *q*. ”

**EXERCISE XXVI**

To appreciate the complexities of the little word “only,” it is useful to notice that it fits at every point in the sentence “I hit him in the eye”:

- Only I hit him in the eye.
- I only hit him in the eye.
- I hit only him in the eye.
- I hit him only in the eye.
- I hit him in only the eye.
- I hit him in the only eye.
- I hit him in the eye only.

Explain what each of these sentences means.

We can next look at sentences of the form “ $p$  if and only if  $q$ ”—so-called biconditionals. If I say that I will clean the barn if and only if Hazel will help me, then I am saying that I will clean it if she helps and I will not clean it if she does not. Translated, this becomes:

$$(H \supset B) \ \& \ (\sim H \supset \sim B)$$

This is equivalent to:

$$(H \supset B) \ \& \ (B \supset H)$$

We thus have an implication going both ways—the characteristic form of a biconditional. In fact, constructions containing the expression “if and only if” do not often appear in everyday speech. They appear almost exclusively in technical or legal writing. In ordinary conversation, we capture the force of a biconditional by saying something like this:

I will clean the barn, but only if Hazel helps me.

The decision whether to translate a remark of everyday conversation into a conditional or a biconditional is often subtle and difficult. We have already noticed that the use of sentences of the form “ $p$  only if  $q$ ” will often conversationally imply a commitment to the biconditional “ $p$  if and only if  $q$ .” In the same way, the use of the conditional “ $p$  if  $q$ ” will often carry this same implication. If I plan to clean the barn whether Hazel helps me or not, it will certainly be misleading—again, a violation of the rule of Quantity—to say that I will clean the barn if Hazel helps me.

We can close this discussion by considering one further, rather difficult case. What is the force of saying “ $p$  unless  $q$ ”? Is this a biconditional, or just a conditional? If it is just a conditional, which way does the implication go? There is a strong temptation to treat this as a biconditional, but the following example shows this to be wrong:

McCain will lose the election unless he carries the South.

This sentence clearly indicates that McCain will lose the election if he does not carry the South. Using abbreviations, we get the following:

$N$  = McCain will carry the South.

$L$  = McCain will lose the election.

$\sim N \supset L$

The original statement does not imply—even conversationally—that McCain will win the election if he does carry the South. Thus,

$p$  unless  $q$  =  $\sim q \supset p$

In short, “unless” means “if not.” We can also note that “ $\sim p$  unless  $q$ ” means the same thing as “ $p$  only if  $q$ ,” and they both are translated thus:

$p \supset q$

Our results can be diagrammed as follows:

	Translates as	Often Conversationally Implies
$p$ if $q$	$q \supset p$	$(p \supset q) \ \& \ (q \supset p)$
$p$ only if $q$	$p \supset q$	$(p \supset q) \ \& \ (q \supset p)$
$p$ unless $q$	$\sim q \supset p$	$(p \supset \sim q) \ \& \ (\sim q \supset p)$

### EXERCISE XXVII

Translate each of the following sentences into symbolic notation, using the suggested symbols as abbreviations.

1. The Reds will win only if the Dodgers collapse. ( $R, D$ )
2. The Steelers will win if their defense holds up. ( $S, D$ )
3. If it rains or snows, the game will be called off. ( $R, S, O$ )
4. If she came home with a trophy and a prize, she must have won the tournament. ( $T, P, W$ )
5. If you order the dinner special, you get dessert and coffee. ( $S, D, C$ )
6. If you order the dinner special, you get dessert; but you can have coffee whether or not you order the dinner special. ( $S, D, C$ )
7. If the house comes up for sale, and if I have the money in hand, I will bid on it. ( $S, M, B$ )
8. If you come to dinner, I will cook you a lobster, if you want me to. ( $D, L, W$ )
9. You can be a success if only you try. ( $S, T$ )
10. You can be a success only if you try. ( $S, T$ )
11. Only if you try can you be a success. ( $S, T$ )

12. You can be a success if you are the only one who tries. (*S, O*)
13. Unless there is a panic, stock prices will continue to rise. (*P, R*)
14. I won't scratch your back unless you scratch mine. (*I, Y*)
15. You will get a good bargain provided you get there early. (*B, E*)
16. You cannot lead a happy life without friends. (Let *H* = You can lead a happy life, and let *F* = You have friends.)
17. The only way that horse will win the race is if every other horse drops dead. (Let *W* = That horse will win the race, and let *D* = Every other horse drops dead.)
18. You should take prescription drugs if, but only if, they are prescribed for you. (*T, P*)
19. The grass will die without rain. (*D, R* = It rains.)
20. Given rain, the grass won't die. (*R, D* = The grass will die.)
21. Unless it doesn't rain, the grass won't die. (*R, D* = The grass will die.)

**EXERCISE XXVIII**

- (a) Translate each of the following arguments into symbolic notation. Then  
 (b) test each argument for truth-functional validity using truth table  
 techniques, and (c) comment on any violations of conversational rules.

Example: Harold is clever; so, if Harold isn't clever, then Anna isn't clever either. (*H, A*)

$$\begin{array}{rcl} \text{(a)} & H & p \\ & \hline & \therefore \sim H \supset \sim A & \therefore \sim p \supset \sim q \end{array}$$

PREMISE			CONCLUSION		
<i>p</i>	<i>q</i>	$\sim p$	$\sim q$	$\sim p \supset \sim q$	
T	T	F	F	T	OK
T	F	F	T	T	OK
F	T	T	F	F	
F	F	T	T	T	

- (c) The argument violates the rule of Relevance, because Anna's cleverness is irrelevant to Harold's cleverness.

1. Jones is brave, so Jones is brave or Jones is brave. (*J*)
2. The Republicans will carry either New Mexico or Arizona; but, since they will carry Arizona, they will not carry New Mexico. (*A, N*)
3. The Democrats will win the election whether they win Idaho or not.  
 Therefore, they will win the election. (*D, I*)

(continued)

4. The Democrats will win the election. Therefore, they will win the election whether they win Idaho or not. ( $D, I$ )
5. The Democrats will win the election. Therefore, they will win the election whether they win a majority or not. ( $D, M$ )
6. If Bobby moves his queen there, he will lose her. Bobby will not lose his queen. Therefore, Bobby will not move his queen there. ( $M, L$ )
7. John will play only if the situation is hopeless. But the situation is hopeless. So John will play. ( $P, H$ )
8. Although Brown will pitch, the Rams will lose. If the Rams lose, their manager will get fired. So their manager will get fired. ( $B, L, F$ )
9. America will win the Olympics unless China does. China will win the Olympics unless Germany does. So America will win the Olympics unless Germany does. ( $A, R, E$ )
10. If you dial 0, you will get the operator. So, if you dial 0 and do not get the operator, then there is something wrong with the telephone. ( $D, O, W$ )
11. The Democrats will run either Jones or Borg. If Borg runs, they will lose the South. If Jones runs, they will lose the North. So the Democrats will lose either the North or the South. ( $J, B, S, N$ )
12. I am going to order either the fish special or the meat special. Either way, I will get soup. So I'll get soup. ( $F, M, S$ )
13. The grass will die if it rains too much or it does not rain enough. If it does not rain enough, it won't rain too much. If it rains too much, then it won't not rain enough. So the grass will die. ( $D = \text{The grass will die}, M = \text{It rains too much}, E = \text{It rains enough}$ )
14. If you flip the switch, then the light will go on. But if the light goes on, then the generator is working. So if you flip the switch, then the generator is working. ( $F, L, G$ ) (This example comes from Charles L. Stevenson.)

## CATEGORICAL LOGIC

*In Chapter 6, we saw how validity can depend on the external connections among propositions. This chapter will demonstrate how validity can depend on the internal structure of propositions. In particular, we will examine two types of categorical arguments—immediate inferences and syllogisms—whose validity or invalidity depends on relations among the subject and predicate terms in their premises and conclusions. Our interest in these kinds of arguments is mostly theoretical. Understanding the theory of the syllogism deepens our understanding of validity, even if this theory is, in some cases, difficult to apply directly to complex arguments in daily life.*

### BEYOND PROPOSITIONAL LOGIC

---

Armed with the techniques developed in Chapter 6, let's look at the following argument:

All squares are rectangles.  
 All rectangles have parallel sides.  
 $\therefore$  All squares have parallel sides.

It is obvious at a glance that the conclusion follows from the premises, so this argument is valid. Furthermore, it seems to be valid in virtue of its form. But it is not yet clear what the form of this argument is. To show the form of this argument, we might try something of the following kind:

$$\begin{array}{c} p \supset q \\ q \supset r \\ \hline \therefore p \supset r \end{array}$$

But this is a mistake—and a bad mistake. We have been using the letters “*p*,” “*q*,” and “*r*” as *propositional variables*—they stand for arbitrary propositions. But the proposition “All squares are rectangles” is not itself composed of two propositions. Nor does it contain “if,” “then” or any other propositional

connective. In fact, if we properly translate the above argument into the language of propositional logic, we get the following result:

$$\begin{array}{c} p \\ q \\ \hline \therefore r \end{array}$$

This, of course, is *not* a valid argument form. But if we look back at the original argument, we see that it is obviously valid. This shows that propositional logic—however adequate it is in its own area—is not capable of explaining the validity of all valid arguments. There is more to logic than propositional logic.

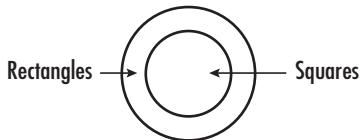
## CATEGORICAL PROPOSITIONS

---

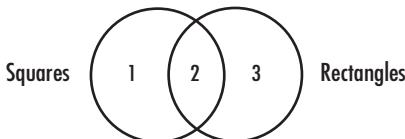
To broaden our understanding of the notion of validity, we will examine a modern version of a branch of logic first developed in ancient times—categorical logic. Categorical logic concerns immediate inferences and syllogisms that are composed of categorical propositions, so we need to begin by explaining what a categorical proposition is.

In the argument above, the first premise asserts some kind of relationship between squares and rectangles; the second premise asserts some kind of relationship between rectangles and things with parallel sides; finally, in virtue of these asserted relationships, the conclusion asserts a relationship between squares and things having parallel sides. Our task is to understand these relationships as clearly as possible so that we can discover the *basis* for the validity of this argument. Again, we shall adopt the strategy of starting from simple cases and then use the insights gained there for dealing with more complicated cases.

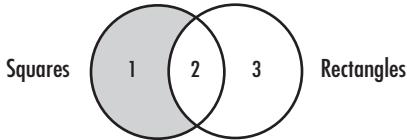
A natural way to represent the relationships expressed by the propositions in an argument is through diagrams. Suppose we draw one circle standing for all things that are squares and another circle standing for all things that are rectangles. The claim that all squares are rectangles may be represented by placing the circle representing squares completely inside the circle representing rectangles.



Another way of representing this relationship is to begin with overlapping circles.

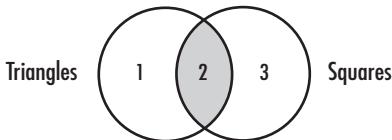


We then shade out the portions of the circles in which nothing exists, according to the proposition we are diagramming. If all squares are rectangles, there is nothing that is a square that is not a rectangle—that is, there is nothing in region 1. So our diagram looks like this:



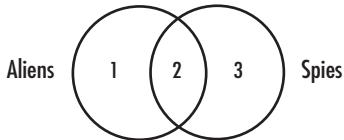
Either method of representation seems plausible. Perhaps the first seems more natural. We shall, however, use the system of overlapping circles, because they will work better when we get to more complex arguments. They are called *Venn diagrams*, after their inventor, John Venn, a nineteenth-century English logician.

Having examined one relationship that can exist between two classes, it is natural to wonder what other relationships might exist. Going to the opposite extreme from our first example, two classes may have *nothing* in common. This relationship could be expressed by saying, “All triangles are not squares,” but it is more common and natural to say, “No triangles are squares.” We diagram this claim by indicating that there is nothing in the overlapping region of things that are both triangles and squares:



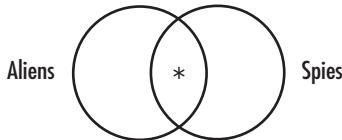
This is one of the relationships that could not be diagrammed by putting one circle inside another. (Just try it!)

In these first two extreme cases, we have indicated that one class is either completely included in another (“All squares are rectangles”) or completely excluded from another (“No triangles are squares”). Sometimes, however, we claim only that two classes have at least *some* things in common. We might say, for example, “Some aliens are spies.” How shall we indicate this relationship in the following diagram?



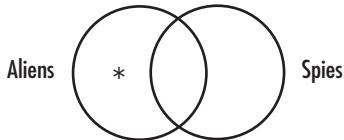
In this case, we do not want to cross out any whole region. We do not want to cross out region 1 because we are not saying that *all* aliens are spies. Plainly, we do not want to cross out region 2, for we are actually saying that some persons *are* both aliens and spies. Finally, we do not want to cross out region 3,

for we are not saying that all spies are aliens. Saying that some aliens are spies does not rule out the possibility that some spies are homegrown. So we need some new device to represent claims that two classes have at least *some* members in common. We shall do this in the following way:

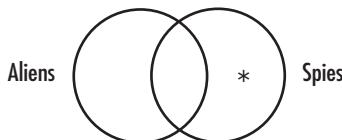


Here the asterisk indicates that there is at least one person who is both an alien and a spy. Notice, by the way, that we are departing a bit from an everyday way of speaking. “Some” is usually taken to mean “more than one”; here we let it mean “at least one.” This makes things simpler and will cause no trouble, so long as we remember that this is what we are using “some” to mean.

Given this new method of diagramming class relationships, we can immediately think of other possibilities. The following diagram indicates that there is someone who is an alien but not a spy. In more natural language, it represents the claim that *some aliens are not spies*.



Next we can indicate that there is someone who is a spy but not an alien. More simply, the claim is that *some spies are not aliens*, and it is represented like this:



These last three claims are, of course, compatible, because there might be some aliens who are spies, some aliens who are not spies, and some spies who are not aliens.

## THE FOUR BASIC CATEGORICAL FORMS

Although two classes can be related in a great many different ways, it is possible to examine many of these relationships in terms of four basic propositional forms:

- |                                |                                    |
|--------------------------------|------------------------------------|
| A: All <i>S</i> is <i>P</i> .  | E: No <i>S</i> is <i>P</i> .       |
| I: Some <i>S</i> is <i>P</i> . | O: Some <i>S</i> is not <i>P</i> . |

These forms are called *categorical forms*, and propositions with these forms are called *categorical propositions*.

As with the propositional forms discussed in the previous chapter, the A, E, I, and O forms for categorical propositions are not themselves propositions, so they are neither true nor false. Instead, they are patterns for whole groups of propositions. We get propositions from these forms by uniformly replacing the variables *S* and *P* with terms that refer to classes of things. For example, “Some spies are not aliens” is a substitution instance of the O propositional form. Nonetheless, we will refer to propositions with the A, E, I, or O form simply as A, E, I, or O propositions, except where this might cause confusion.

A and E propositions are said to be *universal* propositions (because they are about *all S*), and I and O propositions are called *particular* propositions (because they are about *some S*). A and I propositions are described as *affirmative* propositions (because they say what *is P*), and E and O propositions are referred to as *negative* propositions (because they say what *is not P*). Thus, these four basic propositional forms can be described this way:

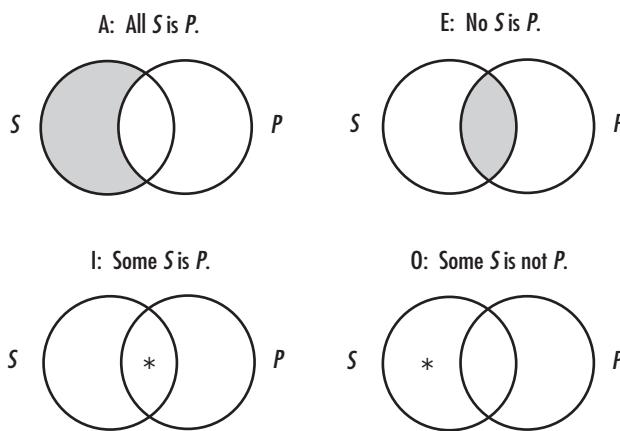
A = Universal Affirmative      E = Universal Negative

I = Particular Affirmative      O = Particular Negative

These four forms fit into the following table:

	<b>Affirmative</b>	<b>Negative</b>
<b>Universal</b>	A: All <i>S</i> is <i>P</i> .	E: No <i>S</i> is <i>P</i> .
<b>Particular</b>	I: Some <i>S</i> is <i>P</i> .	O: Some <i>S</i> is not <i>P</i> .

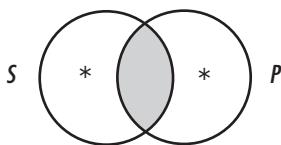
Here are the Venn diagrams for the four basic categorical forms:



These basic categorical forms, together with their labels, classifications, and diagrams, should be memorized, because they will be referred to often in the rest of this chapter.

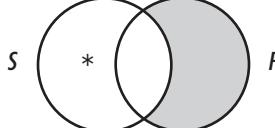
**EXERCISE I**

Using just the four basic categorical forms, indicate what information is given in each of the following diagrams:

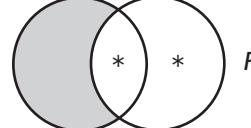
**EXAMPLE:**

No *S* is *P*.  
Some *S* is not *P*.  
Some *P* is not *S*.  
No *P* is *S*.

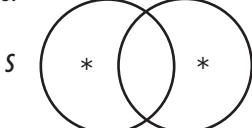
1.



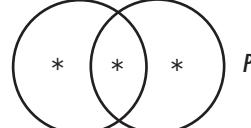
2.



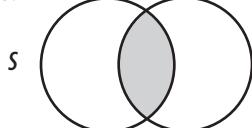
3.



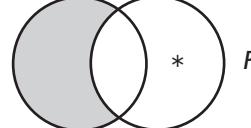
4.



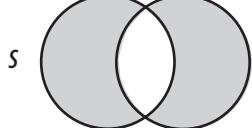
5.



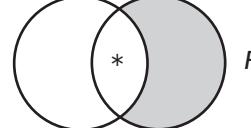
6.



7.



8.

**TRANSLATION INTO THE BASIC CATEGORICAL FORMS**

Propositions with the specific A, E, I, and O forms do not appear often in everyday conversations. Normal people rarely say things like “All whales are mammals. All mammals breathe air. Therefore, all whales breathe air.” Most people talk more like this: “Whales breathe air, since they’re mammals.” Thus,

if our logical apparatus could be applied only to propositions with the explicit forms of A, E, I, and O, then it would apply to few arguments in everyday life.

Fortunately, however, many common statements that are not explicitly in a categorical form can be translated into a categorical form. For example, when someone says, "Whales are mammals," the speaker presumably means to refer to *all* whales, so this statement can be translated into "All whales are mammals," which is an A proposition. We need to be careful, however. If someone says, "Whales are found in the North Atlantic," the speaker probably does *not* mean to refer to *all* whales, because there are many whales in the Pacific as well. Similarly, if someone says, "A whale is a mammal," this can usually be translated as "All whales are mammals," which is an A proposition, but this translation would be inappropriate for "A whale is stranded on the beach," which seems to mean "One whale is stranded on the beach." Thus, we can be misled badly if we look only at the surface structure of what people say. We also need to pay attention to the context when we translate everyday talk into the basic categorical forms.

Despite these complications, it is possible to give some rough-and-ready guides to help in translating many common forms of expression into propositions with the A, E, I, and O forms. Let's begin with one problem that arises for all these categorical forms: They all require a class of things as a predicate. Thus, "All whales are big" and "No whales live on land" should strictly be reformulated as "All whales are big things" and "No whales are things that live on land" or "No whales are land dwellers." This much is easy.

Things get more complicated when we look at the word "all" in A propositions. We have already seen that the word "all" is sometimes dropped in everyday conversation, as in "Whales are mammals." The word "all" can also be moved away from the start of a sentence. "Democrats are all liberal" usually means "All Democrats are liberal," which is an A proposition. Moreover, other words can be used in place of "all." Each of the following claims can, in standard contexts, be translated into an A proposition with the form "All S is P":

Every Republican is conservative.

Any investment is risky.

Anyone who is human is mortal.

Each ant is precious to its mother.

To translate such claims, we sometimes need to construct noun phrases out of adjectives and verbs. These transformations are often straightforward, but sometimes they require ingenuity, and even then they can seem somewhat contorted. For example, both "Only a fool would bungee jump" and "Nobody but a fool would bungee jump" can usually be translated into "All people who bungee jump are fools." This translation might not seem as natural as the originals, but, since the translation has the A form, it explicitly shows that this claim has the logical properties shared by other A propositions.

With some stretching, it is also possible to translate statements about individuals into categorical form. The standard method is to translate "Socrates

is a man" as the A proposition "All things that are Socrates are men." Similarly, "The cannon is about to go off" in a typical context must not be translated as the I proposition "Some cannon is about to go off," because the original statement is about a particular cannon. Instead, the original statement should be translated as the A proposition "All things that are that cannon are about to go off." These translations might seem stilted, but they are necessary in order to apply syllogistic logic to everyday forms of expression.

Similar difficulties arise with the other basic propositional forms. If a woman says, "I am looking for a man who is not attached," and a friend responds, "All of the men in my church are not attached," then this response should probably be translated as "No men in my church are attached," which is an E proposition. In contrast, "All ocean dwellers are not fish" should usually be translated not as the E proposition "No ocean dwellers are fish" but rather as "Not all ocean dwellers are fish." This means "Some ocean dwellers are not fish," which is an O proposition. Thus, some statements with the form "All *S* are not *P*" should be translated as E propositions, but others should be translated as O propositions. (This ambiguity in the form "All *S* are not *P*" explains why it is standard to give E propositions in the less ambiguous form "No *S* is *P*.) Other sentences should also be translated as E propositions even though they do not explicitly contain the word "no." "Underground cables are not easy to repair" and "If a cable is underground, it is not easy to repair" and "There aren't any underground cables that are easy to repair" can all be translated as the E proposition "No underground cables are easy to repair."

Similar complications also arise for I and O propositions. We already saw that "Whales are found in the North Atlantic" should be translated as the I proposition "Some whales are found in the North Atlantic." In addition, some common forms of expression can be translated as O propositions even though they do not contain either the word "not" or the word "some." For example, "There are desserts without chocolate" can be translated as "Some desserts are not chocolate," which is an O proposition.

Because of such complications, there is no mechanical procedure for translating common English sentences into A, E, I, and O propositions. To find the correct translation, you need to think carefully about the sentence and its context.

### EXERCISE II

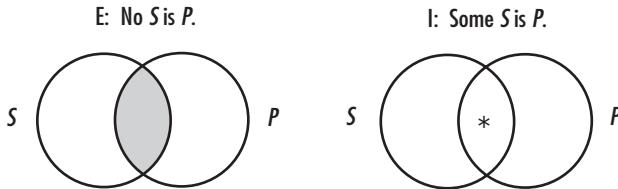
Translate each of the following sentences into an A, E, I, or O proposition. Be sure that the subjects and predicates in your translations use nouns that refer to classes of things (rather than adjectives or verbs). If the sentence can be translated into different forms in different contexts, give each translation and specify a context in which it seems natural.

- |   |   |
|---|---|
| 1. Real men eat ants.<br>2. Bats are not birds. | 3. The hippo is charging.<br>4. The hippo is a noble beast. |
|---|---|

5. Not all crabs live in water.
6. All crabs do not live in water.
7. Movie stars are all rich.
8. If anybody hits me, I will hate them.
9. If anything is broken, it does not work.
10. Somebody loves you.
11. Somebody does not love you.
12. Nobody loves me but my mother.
13. Anybody who is Mormon believes in God.
14. My friends are the only ones who care.
15. Only seniors may take this course.
16. Our pit bull is a good pet.
17. Everything that is cheap is no good.
18. Some things that are expensive are no good.
19. Some things that are cheap are good.
20. Some things that are not cheap are good.
21. Some things that are cheap are not good.
22. Some things that are not cheap are not good.
23. Not all cars have four wheels.
24. There are couples without children.
25. There are no people who hate chocolate.
26. There are people who hate chocolate.
27. Nothing that is purple is an apple.
28. Nothing that is not white is snow.
29. There aren't any runners who are slow.
30. Flamingos aren't friendly.

## CONTRADICTIONES

Once we understand A, E, I, and O propositions by themselves, the next step is to ask how they are related to each other. From their diagrams, some relationships are immediately evident. Consider the Venn diagrams for the E and I propositional forms:

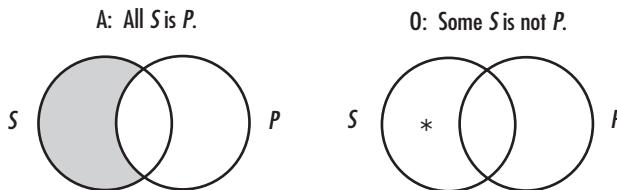


The first diagram has shading in the very same region that contains an asterisk in the second diagram. This makes it obvious that an E proposition and the corresponding I proposition (that is, the I proposition that has the same subject and predicate terms as the E proposition) cannot both be true. For an E proposition to be true, there must be *nothing* in the central region.

But for the corresponding I proposition to be true, there must be *something* in the central region. Thus, they cannot both be true. They also cannot both be false. The only way for an E proposition to be false is for there to be something in the central region, but then the corresponding I proposition is not false but true. The only way for the I proposition to be false is if there is nothing in the central region, and then the E proposition is not false but true. Thus, they cannot both be true, and they cannot both be false. In other words, they always have opposite truth values. This relation is described by saying that these propositions are *contradictories*.

More generally, we can produce a diagram for the denial of a proposition by a simple procedure. The only information given in a Venn diagram is represented either by shading out some region, thereby indicating that nothing exists in it, or by putting an asterisk in a region, thereby indicating that something does exist in it. We are given no information about regions that are unmarked. To represent the denial of a proposition, we simply reverse the information in the diagram. That is, where there is an asterisk, we put in shading; where there is shading, we put in an asterisk. Everything else is left unchanged. Thus, we can see at once that corresponding E and I propositions are denials of one another, so they must always have opposite truth values. This makes them contradictories.

The same relation exists between an A proposition and its corresponding O proposition. Consider their forms:



The diagram for an A proposition has shading exactly where the corresponding O proposition has an asterisk, and they contain no other information. Consequently, corresponding A and O propositions cannot both be false and cannot both be true, so they are contradictories.

### EXERCISE III

1. Is an A proposition a contradictory of its corresponding E proposition?  
Why or why not?
2. Is an I proposition a contradictory of its corresponding O proposition?  
Why or why not?
3. If one proposition is the contradictory of another, is the latter always the contradictory of the former? Why or why not?

## EXISTENTIAL COMMITMENT

It might also seem that an A proposition (with the form “All *S* is *P*”) implies the corresponding I proposition (with the form “Some *S* is *P*”). This, however, raises a difficult problem that logicians have not fully settled. Usually when we make a statement, we are talking about certain specific things. If someone claims that all whales are mammals, that person is talking about whales and mammals and stating a relationship between them. In making this statement, the person seems to be taking the *existence* of whales and mammals for granted. The remark seems to involve what logicians call *existential commitment* to the things referred to in the subject and predicate terms. In the same way, stating an E proposition often seems to commit the speaker to the existence of things in the subject and predicate classes and, thus, to imply an O proposition. For example, someone who says, “No whales are fish” seems committed to “Some whales are not fish.”

In other contexts, however, we seem to use universal (A and E) propositions without committing ourselves to the existence of the things referred to in the subject and predicate terms. For example, if we say, “All trespassers will be fined,” we are not committing ourselves to the existence of any trespassers or to any actual fines for trespassing; we are only saying, “If there are trespassers, then they will be fined.” Similarly, if we tell a sleepy child, “No ghosts are under your bed,” we are not committing ourselves to the existence of ghosts or anything under the bed. Finally, when Newton said, “All bodies that are acted on by no forces are at rest,” he did not commit himself to the existence of bodies that are acted on by no forces. Given these examples of A and E propositions that carry no commitment to the things referred to, it is easy to think of many others.

The question then arises whether we should include existential commitment in our treatment of universal propositions or not. Once more, we must make a decision. (Remember that we had to make decisions concerning the truth-table definitions of both disjunction and conditionals in Chapter 6.) *Classical* logic was developed on the assumption that universal (A and E) propositions carry existential commitment. *Modern* logic makes the opposite decision, treating the claim “All men are mortal” as equivalent to “If someone is a man, then that person is mortal,” and the claim “No men are islands” as equivalent to “If someone is a man, then that person is not an island.” This way of speaking carries no commitment to the existence of any men.

Which approach should we adopt? The modern approach is simpler and has proved more powerful in the long run. For these reasons, we will adopt the modern approach and *not* assign existential commitment to universal (A and E) propositions, so these propositions do not imply particular (I and O) propositions. All the same, there is something beautiful about the classical approach, it has a long and celebrated history, and it does seem appropriate in some contexts, so it is worth exploring in its own right. Still, for the sake of simplicity, we will not develop the classical theory here.

**EXERCISE IV**

Give two new examples of contexts in which:

1. Stating an A proposition does not seem to commit the speaker to the existence of the things to which the subject term refers.
2. Stating an A proposition does not seem to commit the speaker to the existence of the things to which the predicate term refers.
3. Stating an E proposition does not seem to commit the speaker to the existence of the things to which the subject term refers.
4. Stating an E proposition does not seem to commit the speaker to the existence of the things to which the predicate term refers.

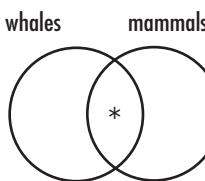
**VALIDITY FOR CATEGORICAL ARGUMENTS**

We have introduced Venn diagrams because they provide an efficient and illuminating way to test the validity of arguments made up of categorical (A, E, I, and O) propositions. The basic idea is simple: An argument made up of categorical propositions is valid if all the information contained in the Venn diagram for the conclusion is already contained in the Venn diagram for the premises. There are only two ways to put information into a Venn diagram: We can either shade out an area or put an asterisk in an area. Hence, to test the validity of an argument made up of categorical propositions, we need only examine the diagram of the conclusion for its information (its shading or asterisks) and then check to see if the diagram for the premises contains this information (the same shading or asterisks).

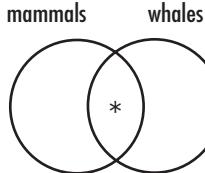
The following simple example will give a general idea of how this works:

*Argument*

Some whales are mammals.

*Diagrams*

$\therefore$  Some mammals are whales.



Notice that the only information contained in the diagram for the conclusion is the asterisk in the overlap between the two circles, and that information is already included in the diagram for the premise. Thus, the argument is valid.

The same method can be used to test argument *forms* for validity. The form of the previous argument and the corresponding diagrams look like this:

<i>Argument Form</i>	<i>Diagrams</i>
Some <i>S</i> are <i>P</i> .	
∴ Some <i>P</i> are <i>S</i> .	

This argument form is valid, because all the information contained in the Venn diagram for the conclusion is contained in the Venn diagram for the premise. And any argument that is a substitution instance of a valid argument form is valid.

Notice that we did not say that an argument is *invalid* if it fails these tests—that is, if some of the information in the Venn diagram for the conclusion (or its form) is not contained in the Venn diagram for the premises (or their forms). As with truth tables in propositional logic (see Chapter 6), Venn diagrams test whether arguments are valid by virtue of a certain form, but some arguments will be valid on a different basis, even though they are not valid by virtue of their categorical form. Here is one example:

<i>Argument</i>	<i>Diagrams</i>
All fathers are male parents.	
∴ All male parents are fathers.	

The Venn diagram for the conclusion includes shading in the circle for male parents, whereas the Venn diagram for the premise includes shading in the circle for fathers, so the premise does not contain the information for the

conclusion. Thus, this *form* of argument is *not* valid, and some arguments of this form are not valid. Nonetheless, this particular argument *is* clearly valid, since it is not possible for the premise to be true when the conclusion is false, for the simple reason that the conclusion cannot be false. Because of such cases, Venn diagrams can show us that an argument is valid, but they cannot prove that an argument is invalid.

Despite this limitation, the method of Venn diagrams can be used to test many different kinds of arguments and argument forms for validity. We will show how this method works for two main kinds of argument: immediate inferences and syllogisms.

## CATEGORICAL IMMEDIATE INFERENCESES

A categorical *immediate inference* is an argument with the following features:

1. It has a single premise. (That is why the inference is called *immediate*.)
2. It is constructed from A, E, I, and O propositions. (That is why the inference is called *categorical*.)

These arguments deserve attention because they occur quite often in everyday reasoning.

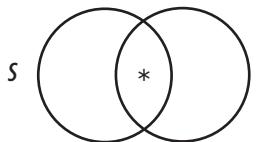
We will focus on the simplest kind of immediate inference, which is *conversion*. We *convert* a proposition (and produce its *converse*) simply by reversing the subject term and the predicate term. By the *subject term*, we mean the term that occurs as the grammatical subject; by the *predicate term*, we mean the term that occurs as the grammatical predicate. In the A proposition “All spies are aliens,” “spies” is the subject term and “aliens” is the predicate term; the converse is “All aliens are spies.”

In this case, identifying the predicate term is straightforward because the grammatical predicate is a noun—a predicate nominative. Often, however, we have to change the grammatical predicate from an adjective to a noun phrase in order to get a noun that refers to a class of things. “All spies are dangerous” becomes “All spies are dangerous things.” Here “spies” is the subject term and “dangerous things” is the predicate term. Although this change is a bit artificial, it is necessary because, when we convert a proposition (that is, reverse its subject and predicate terms), we need a noun phrase to take the place of the grammatical subject. In English we cannot say, “All dangerous are spies,” but we can say, “All dangerous things are spies.”

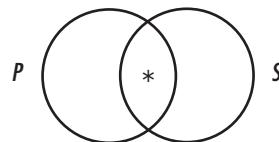
Having explained what conversion is, we now want to know when this operation yields a *valid* immediate inference. To answer this question, we use Venn diagrams to examine the relationship between each of the four basic categorical propositional forms and its converse. The immediate inference is valid if the information contained in the conclusion is also contained in the premise—that is, if any region that is shaded in the conclusion is shaded in the premise, and if any region that contains an asterisk in the conclusion contains an asterisk in the premise.

Two cases are obvious: Both I and E propositions validly convert. From an I proposition with the form “Some  $S$  is  $P$ ,” we may validly infer its converse, which has the form “Some  $P$  is  $S$ .”

I: Some  $S$  is  $P$ .

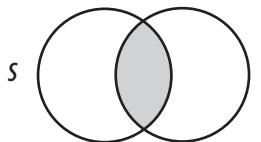


Converse of I: Some  $P$  is  $S$ .

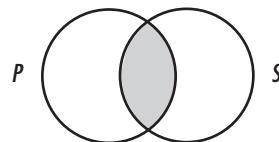


From an E proposition with the form “No  $S$  is  $P$ ,” we may validly infer its converse, which has the form “No  $P$  is  $S$ .”

E: No  $S$  is  $P$ .



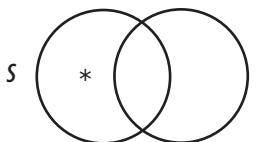
Converse of E: No  $P$  is  $S$ .



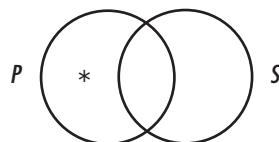
Notice that in both these cases, the information (the asterisk or shading) is in the center of the original diagram, and the diagram for the converse flips the original diagram. Thus, the two diagrams contain the same information, since the diagram for the converse has exactly the same markings in the same areas as does the diagram for the original propositional form. This shows that E and I propositions not only logically *imply* their converses but are also logically *implied by* them. Because the implication runs both ways, these propositions are said to be *logically equivalent* to their converses, and they always have the same truth values as their converses.

The use of a Venn diagram also shows that an O proposition cannot always be converted validly. From a proposition with the form “Some  $S$  is not  $P$ ,” we may not always infer its converse, which has the form “Some  $P$  is not  $S$ .”

O: Some  $S$  is not  $P$ .

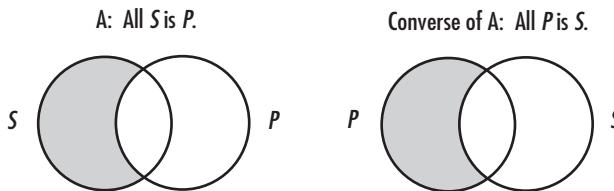


Converse of O: Some  $P$  is not  $S$ .



Notice that in this case the information is not in the center but is instead off to one side. As a result, the information changes when the diagram is flipped. The asterisk is in a different circle—it is in the circle for  $S$  in the diagram for an O proposition, but it is in the circle for  $P$  in the diagram for the converse of the O proposition. That shows that an argument from an O proposition to its converse is not always valid.<sup>1</sup>

Finally, we can see that A propositions also do not always validly convert. From a proposition with the form “All S is P,” we may not always infer its converse, which has the form “All P is S.”



Since the diagram is not symmetrical, the information changes when the diagram is flipped; the shading ends up in a different circle. That shows why this form of argument is not always valid.

Traditionally, other immediate inferences have also been studied, but we will not run through them all here. The single example of conversion is enough to illustrate how Venn diagrams can be used to test some arguments for validity.

#### EXERCISE V

Use Venn diagrams to determine whether the following immediate inferences are valid:

1. All dinosaurs are animals. Therefore, all animals are dinosaurs.
2. Some pterodactyls can fly. Therefore, some flying things are pterodactyls.
3. Some eryopses are not meat eaters. Therefore, some things that eat meat are not eryopses.
4. No tyrannosaurus is a king. Therefore, no king is a tyrannosaurus.
5. Some dinosaurs are reptiles. Therefore, all dinosaurs are reptiles.
6. Some dinosaurs are not alive today. Therefore, no dinosaurs are alive today.
7. All dimetrodons eat meat. Therefore, some dimetrodons eat meat.
8. No dinosaurs are warm-blooded. Therefore, some dinosaurs are not warm-blooded.

### THE THEORY OF THE SYLLOGISM

In an immediate inference, we draw a conclusion directly from a single A, E, I, or O proposition. Moreover, when two categorical propositions are contradictories, the falsity of one can be validly inferred from the truth of the other, and the truth of one can be validly inferred from the falsity of the other. All these forms of argument contain only one premise. The next step in understanding categorical propositions is to consider arguments with two premises.

An important group of such arguments is called *categorical syllogisms*. The basic idea behind these arguments is commonsensical. Suppose you wish to prove that all squares have four sides. A proof should present some *link* or

*connection* between squares and four-sided figures. This link can be provided by some intermediate class, such as rectangles. You can then argue that, because the set of squares is a subset of the set of rectangles and rectangles are a subset of four-sided figures, squares must also be a subset of four-sided figures.

Of course, there are many other ways to link two terms by means of a third term. All such arguments with categorical propositions are called categorical syllogisms. More precisely, a categorical syllogism is any argument such that:

1. The argument has exactly two premises and one conclusion;
2. The argument contains only basic A, E, I, and O propositions;
3. Exactly one premise contains the predicate term;
4. Exactly one premise contains the subject term; and
5. Each premise contains the middle term.

The *predicate term* is simply the term in the predicate of the conclusion. It is also called the *major term*, and the premise that contains the predicate term is called the *major premise*. The *subject term* is the term in the subject of the conclusion. It is called the *minor term*, and the premise that contains the subject term is called the *minor premise*. It is traditional to state the major premise first, the minor premise second.

Our first example of a categorical syllogism then looks like this:

All rectangles are things with four sides.	(Major premise)
<u>All squares are rectangles.</u>	(Minor premise)
∴ All squares are things with four sides.	(Conclusion)

Subject term = "Squares"

Predicate term = "Things with four sides"

Middle term = "Rectangles"

To get the form of this syllogism, we replace the terms with variables:

All M is P.	
<u>All S is M.</u>	
∴ All S is P.	

Of course, many other arguments fit the definition of a categorical syllogism. Here is one with a negative premise:

No ellipses are things with sides.	
<u>All circles are ellipses.</u>	
∴ No circles are things with sides.	

The next categorical syllogism has a particular premise:

All squares are things with equal sides.	
<u>Some squares are rectangles.</u>	
∴ Some rectangles are things with equal sides.	

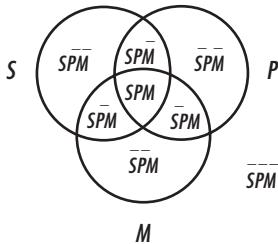
**EXERCISE VI**

In each of the last two syllogisms, what is the subject term? The predicate term? The middle term? The major premise? The minor premise? The form of the syllogism (using *S*, *P*, and *M*)? Is the syllogism valid? Why or why not?

**HONORS EXERCISE**

Given the restrictions in the definition of a categorical syllogism, there are exactly 256 possible forms of categorical syllogism. Explain why.

*VENN DIAGRAMS FOR SYLLOGISMS.* In a previous section, we used Venn diagrams to test the validity of immediate inferences. Immediate inferences contain only two terms or classes, so the corresponding Venn diagrams need only two overlapping circles. Categorical syllogisms contain three terms or classes. To reflect this, we will use diagrams with three overlapping circles. If we use a bar over a letter to indicate that things in the area are not in the class (so that  $\bar{S}$  indicates what is not in *S*), then our diagram looks like this:



This diagram has eight different areas, which can be listed in an order that resembles a truth table:

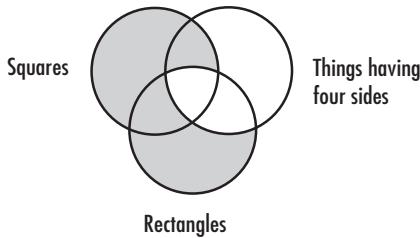
<i>S</i>	<i>P</i>	<i>M</i>
<i>S</i>	$\bar{P}$	$\bar{M}$
$\bar{S}$	$\bar{P}$	<i>M</i>
$\bar{S}$	$\bar{P}$	$\bar{M}$
$\bar{S}$	<i>P</i>	<i>M</i>
$\bar{S}$	<i>P</i>	$\bar{M}$
$\bar{S}$	$\bar{P}$	<i>M</i>
$\bar{S}$	$\bar{P}$	$\bar{M}$

Notice that, if something is neither an *S* nor a *P* nor an *M*, then it falls completely outside the system of overlapping circles. In every other case, a thing is assigned to one of the seven compartments within the system of overlapping circles.

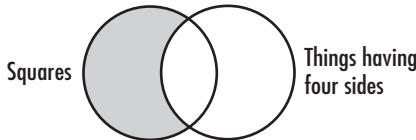
**TESTING SYLLOGISMS FOR VALIDITY.** To test the validity of a syllogism using a Venn diagram, we first fill in the diagram to indicate the information contained in the premises. Remember that the only information contained in a Venn diagram is indicated either by shading out an area or by putting an asterisk in it. The argument is valid if the information expressed by the conclusion is already contained in the diagram for the premises.<sup>2</sup> To see this, consider the diagrams for examples that we have already given:

- All rectangles have four sides.  
All squares are rectangles.  
 ∴ All squares have four sides.

Here's the diagram for the premises:



Here's the diagram for the conclusion:

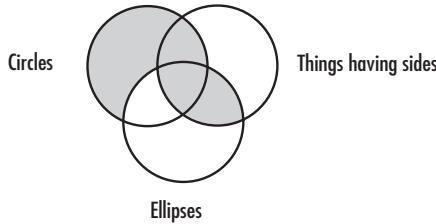


This diagram for the conclusion contains only the information that nothing is in the circle for squares that is not also in the circle for things having four sides. In the diagram for the premises, all the things that are squares are corralled into the region of things that have four sides. Thus, the diagram for the premises contains all of the information in the diagram for the conclusion. That shows that this syllogism is valid.

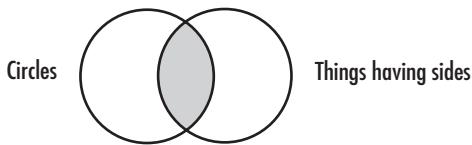
Next, let's try a syllogism with a negative premise:

- No ellipses have sides.  
All circles are ellipses.  
 ∴ No circles have sides.

Here's the diagram for the premises:



We diagram the conclusion “No circles have sides” as follows:



That information is clearly already contained in the Venn diagram for the premises, so this syllogism is also valid.

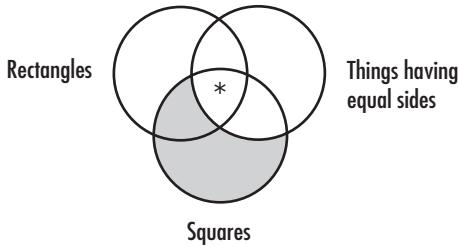
Let's try a syllogism with a particular premise:

All squares have equal sides.

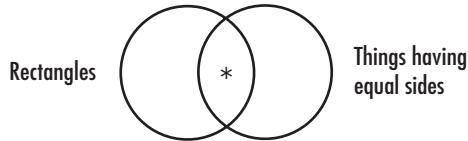
Some squares are rectangles.

∴ Some rectangles have equal sides.

It is a good strategy to diagram a universal premise *before* diagramming a particular premise. The diagram for the above argument then looks like this:



Here's the diagram for the conclusion—that there is something that is a rectangle that has equal sides:



The asterisk in the middle area of this diagram says that something is in both circles, and that information already appears in the diagram for the premises, so this argument is valid.

So far we have looked only at valid syllogisms. Let's see how this method applies to invalid syllogisms. Here is one:

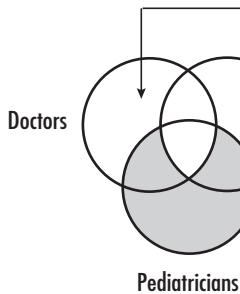
All pediatricians are doctors.

All pediatricians like children.

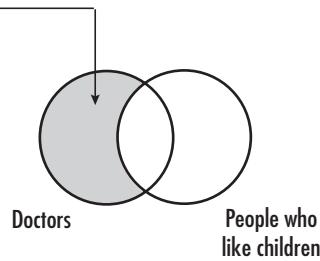
∴ All doctors like children.

We can diagram the premises at the left and the conclusion at the right:

Premises:  
All pediatricians are doctors.  
All pediatricians like children.



Conclusion:  
All doctors like children.



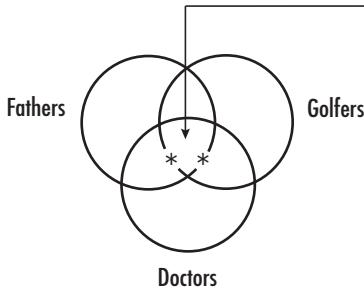
It is evident that the information in the diagram for the conclusion is *not* already contained in the diagram for the premises. The arrow shows differences in informational content. Thus, this form of syllogism is not valid.

Notice that the difference between these diagrams not only tells us *that* this form of syllogism is invalid; it also tells us *why* it is invalid. In the diagram for the premises, there is no shading in the upper left area, which includes people who are doctors but are not pediatricians and do not like children. This shows that the premises do not rule out the possibility that some people are doctors without being pediatricians or liking children. But if anyone is a doctor and not a person who likes children, then it is not true that all doctors like children. Because this is the conclusion of the syllogism, the premises do not rule out all of the ways in which the conclusion might be false. As a result, this conclusion does not follow by virtue of categorical form.<sup>3</sup>

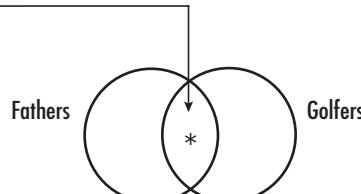
Here is an example of an invalid syllogism with particular premises:

- Some doctors are golfers.
- Some fathers are doctors.
- ∴ Some fathers are golfers.

Premises:



Conclusion:



Examine this diagram closely. Notice that in diagramming “Some doctors are golfers,” we had to put an asterisk *on the boundary* of the circle for

fathers, because we were not given information saying whether anything falls into the category of fathers or not. For the same reason, we had to put an asterisk on the boundary of the circle for golfers when diagramming "Some fathers are doctors." The upshot was that we did not indicate that anything exists in the region of overlap between fathers and golfers. But this is what the conclusion demands, so the form of this syllogism is not valid.

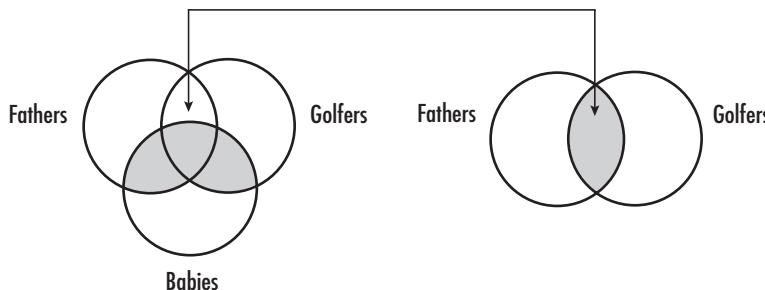
Here is an invalid syllogism with negative premises:

No babies are golfers.

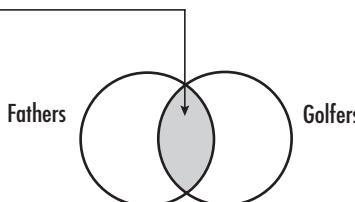
No fathers are babies.

∴ No fathers are golfers.

Premises:



Conclusion:



Again, we see that the form of this syllogism is not valid, because the entire area of overlap between the circles is shaded in the diagram for the conclusion, but part of that area is not shaded in the diagram for the premises.

The method of Venn diagrams is adequate for deciding the validity or invalidity of all possible forms of categorical syllogism. To master this method, all you need is a little practice.

### EXERCISE VII

Using Venn diagrams, test the following syllogistic forms for validity:

- |  |  |
|--|--|
| 1. All <i>M</i> is <i>P</i> .<br>All <i>M</i> is <i>S</i> .  | 4. All <i>P</i> is <i>M</i> .<br>Some <i>M</i> is <i>S</i> .     |
| <u>∴ All <i>S</i> is <i>P</i>.</u>                           | <u>∴ Some <i>S</i> is <i>P</i>.</u>                              |
| 2. All <i>P</i> is <i>M</i> .<br>All <i>M</i> is <i>S</i> .  | 5. All <i>P</i> is <i>M</i> .<br>Some <i>S</i> is <i>M</i> .     |
| <u>∴ All <i>S</i> is <i>P</i>.</u>                           | <u>∴ Some <i>S</i> is <i>P</i>.</u>                              |
| 3. All <i>M</i> is <i>P</i> .<br>Some <i>M</i> is <i>S</i> . | 6. All <i>P</i> is <i>M</i> .<br>Some <i>S</i> is not <i>M</i> . |
| <u>∴ Some <i>S</i> is <i>P</i>.</u>                          | <u>∴ Some <i>S</i> is not <i>P</i>.</u>                          |

- |  |  |
|--|--|
| <p>7. All <math>M</math> is <math>P</math>.<br/> <u>Some <math>S</math> is not <math>M</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math>.</p> <p>8. All <math>M</math> is <math>P</math>.<br/> <u>Some <math>M</math> is not <math>S</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math>.</p> <p>9. No <math>M</math> is <math>P</math>.<br/> <u>Some <math>S</math> is <math>M</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math>.</p> <p>10. No <math>P</math> is <math>M</math>.<br/> <u>Some <math>S</math> is <math>M</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math>.</p> <p>11. No <math>P</math> is <math>M</math>.<br/> <u>Some <math>S</math> is not <math>M</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math>.</p> <p>12. No <math>M</math> is <math>P</math>.<br/> <u>Some <math>S</math> is not <math>M</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math></p> <p>13. No <math>P</math> is <math>M</math>.<br/> <u>Some <math>M</math> is not <math>S</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math>.</p> | <p>14. No <math>P</math> is <math>M</math>.<br/> <u>No <math>M</math> is <math>S</math>.</u><br/> <math>\therefore</math> No <math>S</math> is <math>P</math>.</p> <p>15. No <math>P</math> is <math>M</math>.<br/> <u>All <math>M</math> is <math>S</math>.</u><br/> <math>\therefore</math> No <math>S</math> is <math>P</math>.</p> <p>16. No <math>P</math> is <math>M</math>.<br/> <u>All <math>S</math> is <math>M</math>.</u><br/> <math>\therefore</math> No <math>S</math> is <math>P</math>.</p> <p>17. All <math>P</math> is <math>M</math>.<br/> <u>No <math>S</math> is <math>M</math>.</u><br/> <math>\therefore</math> No <math>S</math> is <math>P</math>.</p> <p>18. All <math>M</math> is <math>P</math>.<br/> <u>No <math>S</math> is <math>M</math>.</u><br/> <math>\therefore</math> No <math>S</math> is <math>P</math>.</p> <p>19. Some <math>M</math> is <math>P</math>.<br/> <u>Some <math>M</math> is not <math>S</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is not <math>P</math>.</p> <p>20. Some <math>P</math> is <math>M</math>.<br/> <u>Some <math>S</math> is not <math>M</math>.</u><br/> <math>\therefore</math> Some <math>S</math> is <math>P</math>.</p> |
|--|--|

### EXERCISE VIII

Explain why it is a good strategy to diagram a universal premise before diagramming a particular premise in a syllogism with both.

**PROBLEMS IN APPLYING THE THEORY OF THE SYLLOGISM.** After mastering the techniques for evaluating syllogisms, students naturally turn to arguments that arise in daily life and attempt to use these newly acquired skills. They are often disappointed with the results. The formal theory of the syllogism seems to bear little relationship to everyday arguments, and there does not seem to be any easy way to bridge the gap.

This gap between formal theory and its application occurs for a number of reasons. First, as we saw in Chapters 2 and 5, our everyday discourse leaves much unstated. Many things are conversationally implied rather

than explicitly asserted. We do not feel called on to say many things that are matters of common agreement. Before we can apply the theory of the syllogism to everyday arguments, these things that are simply understood must be made explicit. This is often illuminating, and sometimes boring, but it usually involves a great deal of work. Second, the theory of the syllogism applies to statements only in a highly stylized form. Before we apply the theory of the syllogism to an argument, we must cast its premises and conclusion into the basic A, E, I, and O forms. As we saw earlier in this chapter, the needed translation is not always simple or obvious. It may not always be possible. For these and related reasons, modern logicians have largely abandoned the project of reducing all reasoning to syllogisms.

Why study the theory of the syllogism at all, if it is hard to apply in some circumstances and perhaps impossible to apply in others? The answer to this question was given at the beginning of Chapter 6. The study of formal logic is important because it deepens our insight into a central notion of logic: *validity*. Furthermore, the argument forms we have studied do underlie much of our everyday reasoning, but so much else is going on in a normal conversational setting that this dimension is often hidden. By examining arguments in idealized forms, we can study their validity in isolation from all the other factors at work in a rich conversational setting.

There is a difference, then, between the techniques developed in Chapters 1–5 and the techniques developed in Chapters 6 and 7. The first five chapters presented methods of informal analysis that may be applied directly to the rich and complex arguments that arise in everyday life. These methods of analysis are not wholly rigorous, but they do provide practical guides for the analysis and evaluation of actual arguments. The chapters concerning formal logic have the opposite tendency. In comparison with the first five chapters, the level of rigor is very high, but the range of application is correspondingly smaller. In general, the more rigor and precision you insist on, the less you can talk about.

#### DISCUSSION QUESTIONS

1. What are the chief differences between the logical procedures developed in this chapter and those developed in Chapter 6 on propositional logic?
2. If we evaluate arguments as they occur in everyday life by using the exact standards developed in Chapters 6 and 7, we discover that our everyday arguments rarely satisfy these standards, at least explicitly. Does this show that most of our ordinary arguments are illogical? What else might it show?

## NOTES

<sup>1</sup>We say “not always” rather than simply “not,” because there are some strange cases—logicians call them “degenerate cases”—for which inferences of this pattern are valid. For example, from “Some men are not men,” we may validly infer “Some men are not men.” Here, by making the subject term and the predicate term the same, we trivialize conversion. Keeping cases of this kind in mind, we must say that the inference from an O proposition to its converse is usually, but not always, invalid. In contrast, the set of valid arguments holds in all cases, including degenerate cases.

<sup>2</sup>We cannot say “only if” here because of degenerate cases of categorical syllogisms that are valid, but not by virtue of their syllogistic form. Here is one example: “All numbers divisible by two are even. No prime number other than two is divisible by two. Therefore, no prime number other than two is even.” This syllogism is valid because it is not possible that its premises are true and its conclusion is false, but other syllogisms with this same form are not valid.

<sup>3</sup>We need to add “*by virtue of its categorical form*,” because, as we saw above, it still might be valid on some other basis. In this particular example, however, nothing else makes this argument valid.



## HOW TO EVALUATE ARGUMENTS: INDUCTIVE STANDARDS

*Previous chapters have been concerned primarily with deductive arguments that aim at validity. Many arguments encountered in daily life, however, are not intended to meet this standard of validity. They are only supposed to provide reasons (perhaps very strong reasons) for their conclusions. Such arguments are called inductive and will be the focus of Part III. This part begins with a discussion of the nature of inductive standards and arguments followed by a survey of five forms of inductive argument: statistical generalizations, statistical applications, inference to the best explanation, arguments from analogy, and causal reasoning. The next topic is probability, because, as we will see, the inductive standard of strength can be understood in terms of probability. Part III will close by discussing how probabilities get deployed in decision making.*



## ARGUMENTS TO AND FROM GENERALIZATIONS

This chapter begins our investigation of inductive arguments by distinguishing the inductive standard of strength from the deductive standard of validity. Inductive arguments are defined as arguments that are intended to be strong rather than valid. Two common examples of inductive arguments are discussed next. In statistical generalizations, a claim is made about a population on the basis of features of a sample of that population. In statistical applications, a claim is made about members of a population on the basis of features of the population. Statistical generalizations take us up from samples to general claims, and statistical applications then take us back down to individual cases.

### INDUCTION VERSUS DEDUCTION

---

The distinction between deductive arguments and inductive arguments can be drawn in a variety of ways, but the fundamental difference concerns the relationship that is claimed to hold between the premises and the conclusion for each type of argument. An argument is *deductive* insofar as it is intended or claimed to be *valid*. As we know from Chapter 5, an argument is valid if and only if it is impossible for the conclusion to be false when its premises are true. The following is a valid deductive argument:

All ravens are black.

∴ If there is a raven on top of Pikes Peak, then it is black.

Because the premise lays down a universal principle governing all ravens, if it's true, then it *must* be true of all ravens (if any) on top of Pikes Peak. This same relationship does not hold for invalid arguments. Nonetheless, arguments that are not valid can still be deductive if they are intended or claimed to be valid.

In contrast, inductive arguments are not intended to be valid, so they should not be criticized for being invalid. The following is an example of an inductive argument:

All ravens that we have observed so far are black.

∴ All ravens are black.

Here we have drawn an inductive inference from the characteristics of *observed* ravens to the characteristics of *all* ravens, most of which we have not observed. Of course, the premise of this argument *could be* true, yet the conclusion turn out to be false. A raven that has not yet been observed might be albino. The obviousness of this possibility suggests that someone who gives this argument does not put it forth as valid, so it is not a deductive argument. Instead, the premise is put forth as a *reason* or *support* for the conclusion. When an argument is not claimed to be valid but is intended only to provide a reason for the conclusion, the argument is *inductive*.

Because inductive arguments are supposed to provide reasons, and reasons vary in strength, inductive arguments can be evaluated as *strong* or *weak*, depending on the strength of the reasons that they provide for their conclusions. If we have seen only ten ravens, and all of them were in our backyard, then the above argument gives at most a very weak reason to believe that all ravens are black. But, if we have traveled around the world and seen over half the ravens that exist, then the above argument gives a strong reason to believe that all ravens are black. Inductive arguments are usually intended to provide strong support for their conclusions, in which case they can be criticized if the support they provide is not strong enough for the purposes at hand.

The most basic distinction, then, is not between two kinds of argument but is instead between two standards for evaluating arguments. The deductive standard is validity. The inductive standard is strength. Arguments themselves are classified as either deductive or inductive in accordance with the standard that they are intended or claimed to meet.

There are several important differences between deductive and inductive standards. One fundamental feature of the deductive standard of validity is that adding premises to a valid argument cannot make it invalid. The definition of validity guarantees this: In a valid argument, it is not possible for the premises to be true without the conclusion being true as well. If any further premises could change this, then it would be possible for this relationship not to hold, so the argument would not be valid after all. Additional information might, of course, lead us to question the truth of one of the premises, but that is another matter.

The situation is strikingly different when we deal with inductive arguments. To cite a famous example, before the time of Captain Cook's voyage to Australia, Europeans had observed a great many swans, and every one of them was white. Thus, up to that time Europeans had very strong inductive evidence to support the claim that all swans are white. Then Captain Cook discovered black swans in Australia. What happens if we add this new piece of information to the premises of the original inductive argument? Provided that we accept Cook's report, we now produce a sound *deductive* argument in behalf of the opposite claim that *not* all swans are white; for, if some swans are black, then not all of them are white. This, then, is a feature of the inductive standard of strength: No matter how strong an inductive argument is, the possibility remains open that further information can undercut, perhaps completely, the strength of the argument and the support that the premises

give to the conclusion. Because inductive strength and inductive arguments can always be defeated in this way, they are described as *defeasible*. Valid deductive arguments do not face a similar peril, so they are called *indefeasible*.

A second important difference between inductive and deductive standards is that inductive strength comes in degrees, but deductive validity does not. An argument is either valid or invalid. There is no question of how much validity an argument has. In contrast, inductive arguments can be more or less strong. The more varied ravens or swans we observe, the stronger the inductive arguments above. Some inductive arguments are extremely strong and put their conclusions beyond any reasonable doubt. Other inductive arguments are much weaker, even though they still have some force.

Because of the necessary relationship between the premises and the conclusion of a valid deductive argument, it is often said that the premises of valid deductive arguments (if true) provide *conclusive* support for their conclusions, whereas true premises of strong inductive arguments provide only *partial* support for their conclusions. There is something to this. Because the premises of a valid deductive argument necessitate the truth of the conclusion, if those premises are definitely known to be true, then they do supply conclusive reasons for the conclusion. The same cannot be said for inductive arguments.

It would be altogether misleading, however, to conclude from this that inductive arguments are inherently inferior to deductive arguments in supplying a justification or ground for a conclusion. In the first place, inductive arguments often place matters beyond any reasonable doubt. It is possible that the next pot of water will not boil at any temperature, however high, but this is not something we worry about. We do not take precautions against it, and we shouldn't.

More important, deductive arguments normally enjoy no advantages over their inductive counterparts. We can see this by comparing the two following arguments:

DEDUCTIVE

- All ravens are black.  
∴ If there is a raven on top  
of Pikes Peak, it is black.

INDUCTIVE

- All observed ravens are black.  
∴ If there is a raven on top of  
Pikes Peak, it is black.

Of course, it is true for the deductive argument (and not true for the inductive argument) that if the premise is true, then the conclusion must be true. This may seem to give an advantage to the deductive argument over the inductive argument. But before we can decide how much support a deductive argument gives its conclusion, we must ask whether its premises are, after all, true. That is not something we can just take for granted. If we examine the premises of these two arguments, we see that it is easier to establish the truth of the premise of the inductive argument than it is to establish the truth of the premise of the deductive argument. If we have observed carefully and kept good records, then we might be fully confident that all *observed* ravens have been black. On the other hand, how can we show that *all* ravens (observed

and unobserved—past, present, and future) are black? The most obvious way (though there may be other ways) would be to observe ravens to see whether they are black or not. This, of course, involves producing an inductive argument (called a statistical generalization) for the premise of the deductive argument. Here our confidence in the truth of the premise of the deductive argument should be no greater than our confidence in the strength of the inference in the statistical generalization. In this case—and it is not unusual—the deductive argument provides no stronger grounds in support of its conclusion than does its inductive counterpart, because any reservations we might have about the *strength* of the inductive inference will be paralleled by doubts concerning the *truth* of the premise of the deductive argument.

We will also avoid the common mistake of saying that deductive arguments always move from the general to the particular, whereas inductive arguments always move from the particular to the general. In fact, both sorts of arguments can move in either direction. There are inductive arguments intended to establish particular matters of fact, and there are deductive arguments that involve generalizations from particulars. For example, when scientists assemble empirical evidence to determine whether the extinction of the dinosaurs was caused by the impact of a meteor, their discussions are models of inductive reasoning. Yet they are not trying to establish a generalization or a scientific law. Instead, they are trying to determine whether a particular event occurred some 65 million years ago. Inductive reasoning concerning particular matters of fact occurs constantly in everyday life as well, for example, when we check to see whether our television reception is being messed up by someone using a hair dryer. Deductive arguments from the particular to the general also exist, though they tend to be trivial, and hence boring. Here's one:

Benjamin Franklin was the first postmaster general; therefore, anyone who is identical with Benjamin Franklin was the first postmaster general.

Of course, many deductive arguments do move from the general to the particular, and many inductive arguments do move from particular premises to a general conclusion. It is important to remember, however, that this is not the *definitive* difference between these two kinds of arguments. What makes deductive arguments deductive is precisely that they are intended to meet the deductive standard of validity, and what makes inductive arguments inductive is just that they are not intended to be deductively valid but are, instead, intended to be inductively strong.

#### EXERCISE I

Assuming a standard context, label each of the following arguments as deductive or inductive. Explain what it is about the words or form of argument that indicates whether or not each argument is intended or claimed to be valid. If it is not clear whether the argument is inductive or deductive, say why.

1. The sun is coming out, so the rain will probably stop soon.
2. It's going to rain tomorrow, so it will either rain or be clear tomorrow.
3. No woman has ever been elected president. Therefore, no woman will ever be elected president.
4. Diet cola never keeps me awake at night. I know because I drank it just last night without any problems.
5. The house is a mess, so Jeff must be home from college.
6. If Harold were innocent, he would not go into hiding. Since he is hiding, he must not be innocent.
7. Nobody in Paris seems to understand me, so either my French is rotten or Parisians are unfriendly.
8. Because both of our yards are near rivers in Tennessee, and my yard has lots of mosquitoes, there must also be lots of mosquitoes in your yard.
9. Most likely, her new husband speaks English with an accent, because he comes from Germany, and most Germans speak English with an accent.
10. There is no even number smaller than two, so one is not an even number.

## STATISTICAL GENERALIZATIONS

One classic example of an inductive argument is an opinion poll. Suppose a candidate wants to know how popular she is with voters. Because it would be practically impossible to survey all voters, she takes a sample of voting opinion and then infers that the opinions of those sampled indicate the overall opinion of voters. Thus, if 60 percent of the voters sampled say that they will vote for her, she concludes that she will get around 60 percent of the vote in the actual election. As we shall see later, inferences of this kind often go wrong, even when made by experts, but the general pattern of this reasoning is quite clear: Statistical features of a sample are used to make statistical claims about the population as a whole.

Basically the same form of reasoning can be used to reach a universal conclusion. An example is the inductive inference discussed at the start of this chapter: All observed ravens are black, so all ravens are black. Again, we sample part of a population to draw a conclusion about the whole. Arguments of this form, whether the conclusion is universal or partial (as when it cites a particular percentage), are called *statistical generalizations*.

How do we assess such inferences? To begin to answer this question, we can consider a simple example of a statistical generalization. On various occasions, Harold has tried to use Canadian quarters in American vending machines and found that they have not worked. From this he draws the

conclusion that Canadian quarters do not work in American vending machines. Harold's inductive reasoning looks like this:

In the past, when I tried to use Canadian quarters in American vending machines, they did not work.

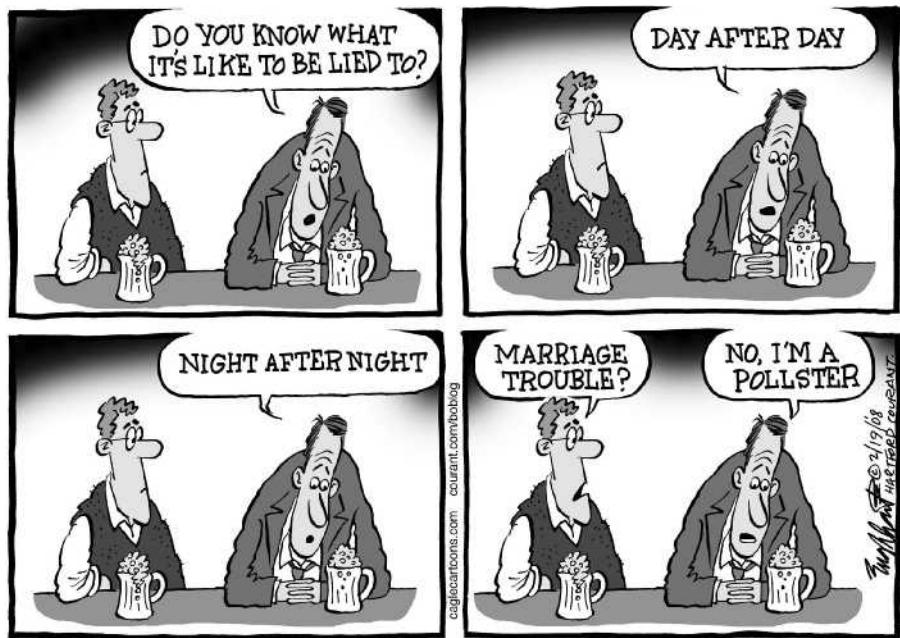
∴ Canadian quarters do not work in American vending machines.

The force of the conclusion is that Canadian quarters *never* work in American vending machines.

In evaluating this argument, what questions should we ask? We can start with a question that we should ask of any argument.

### SHOULD WE ACCEPT THE PREMISES?

Perhaps Harold has a bad memory, has kept bad records, or is a poor observer. For some obscure reason, he may even be lying. It is important to ask this question explicitly, because fairly often the premises, when challenged, will not stand up to scrutiny.



Bob Englehart/Cagle Cartoons, Inc.

If we decide that the premises are acceptable (that is, true and justified), then we can shift our attention to the relationship between the premises and the conclusion and ask how much support the premises give to the conclusion. One commonsense question is this: "How many times has Harold tried to use Canadian quarters in American vending machines?" If the answer is "Once,"

then our confidence in his argument should drop to almost nothing. So, for statistical generalizations, it is always appropriate to ask about the size of the sample.

### IS THE SAMPLE LARGE ENOUGH?

One reason we should be suspicious of small samples is that they can be affected by runs of luck. Suppose Harold flips a Canadian quarter four times and it comes up heads each time. From this, he can hardly conclude that Canadian quarters always come up heads when flipped. He could not even reasonably conclude that *this* Canadian quarter would always come up heads when flipped. The reason for this is obvious enough: If you spend a lot of time flipping coins, runs of four heads in a row are not all that unlikely (the probability is actually one in sixteen), and therefore samples of this size can easily be distorted by chance. On the other hand, if Harold flipped the coin twenty times and it continued to come up heads, he would have strong grounds for saying that this coin, at least, will always come up heads. In fact, he would have strong grounds for thinking that he has a two-headed coin. Because an overly small sample can lead to erroneous conclusions, we need to make sure that our sample includes enough trials.

How many is enough? On the assumption, for the moment, that our sampling has been fair in all other respects, how many samples do we need to provide the basis for a strong inductive argument? This is not always an easy question to answer, and sometimes answering it demands subtle mathematical techniques. Suppose your company is selling 10 million computer chips to the Department of Defense, and you have guaranteed that no more than 0.2 percent of them will be defective. It would be prohibitively expensive to test all the chips, and testing only a dozen would hardly be enough to reasonably guarantee that the total shipment of chips meets the required specifications. Because testing chips is expensive, you want to test as few as possible; but because meeting the specifications is crucial, you want to test enough to guarantee that you have done so. Answering questions of this kind demands sophisticated statistical techniques beyond the scope of this text.

Sometimes, then, it is difficult to decide how many instances are needed to give reasonable support to inductive generalizations; yet many times it is obvious, without going into technical details, that the sample is too small. Drawing an inductive conclusion from a sample that is too small can lead to the fallacy of *hasty generalization*. It is surprising how common this fallacy is. We see a person two or three times and find him cheerful, and we immediately leap to the conclusion that he is a cheerful person. That is, from a few instances of cheerful behavior, we draw a general conclusion about his personality. When we meet him later and find him sad, morose, or grouchy, we then conclude that he has changed—thus swapping one hasty generalization for another.

By making our samples sufficiently large, we can guard against distortions due to “runs of luck,” but even very large samples can give us a poor basis for a statistical generalization. Suppose that Harold has tried hundreds of times to use

a Canadian quarter in an American vending machine, and it has never worked. This will increase our confidence in his generalization, but size of sample alone is not a sufficient ground for a strong inductive argument. Suppose that Harold has tried the same coin in hundreds of American vending machines, or tried a hundred different Canadian coins in the same vending machine. In the first case, there might be something wrong with this particular coin; in the second case, there might be something wrong with this particular vending machine. In neither case would he have good grounds for making the general claim that *no* Canadian quarters work in *any* American vending machine. This leads us to the third question we should ask of any statistical generalization.

### IS THE SAMPLE BIASED?

When the sample, however large, is not representative of the population, then it is said to be unfair or biased. Here we can speak of the fallacy of *biased sampling*.

One of the most famous errors of biased sampling was committed by a magazine named the *Literary Digest*. Before the presidential election of 1936, this magazine sent out 10 million questionnaires asking which candidate the recipient would vote for: Franklin Roosevelt or Alf Landon. It received 2.5 million returns, and on the basis of the results, confidently predicted that Landon would win by a landslide: 56 percent for Landon to only 44 percent for Roosevelt. When the election results came in, Roosevelt had won by an even larger landslide in the opposite direction: 62 percent for Roosevelt to a mere 38 percent for Landon.

What went wrong? The sample was certainly large enough; in fact, by contemporary standards it was much larger than needed. It was the way the sample was selected, not its size, that caused the problem: The sample was randomly drawn from names in telephone books and from club membership lists. In 1936 there were only 11 million payphones in the United States, and many of the poor—especially the rural poor—did not have payphones. During the Great Depression there were more than nine million unemployed in America; they were almost all poor and thus underrepresented on club membership lists. Finally, a large percentage of these underrepresented groups voted for Roosevelt, the Democratic candidate. As a result of these biases in its sampling, along with some others, the *Literary Digest* underestimated Roosevelt's percentage of the vote by a whopping 18 percent.

Looking back, it may be hard to believe that intelligent observers could have done such a ridiculously bad job of sampling opinion, but the story repeats itself, though rarely on the grand scale of the *Literary Digest* fiasco. In 1948, for example, the Gallup poll, which had correctly predicted Roosevelt's victory in 1936, predicted, as did other major polls, a clear victory for Thomas Dewey over Harry Truman. Confidence was so high in this prediction that the *Chicago Tribune* published a banner headline declaring that Dewey had won the election before the votes were actually counted.

What went wrong this time? The answer here is more subtle. The Gallup pollsters (and others) went to great pains to make sure that their sample was representative of the voting population. The interviewers were told to poll a certain number of people from particular social groups—rural poor, suburban middle class, urban middle class, ethnic minorities, and so on—so that the proportions of those interviewed matched, as closely as possible, the proportions of those likely to vote. (The *Literary Digest* went bankrupt after its incorrect prediction, so the pollsters were taking no chances.) Yet somehow bias crept into the sampling; the question was, “How?” One speculation was that a large percentage of those sampled did not tell the truth when they were interviewed; another was that a large number of people changed their minds at the last minute. So perhaps the data collected were not reliable. The explanation generally accepted was more subtle. Although Gallup’s workers were told to interview specific numbers of people from particular classes (so many from the suburbs, for example), they were not instructed to choose people randomly from within each group. Without seriously thinking about it, they tended to go to “nicer” neighborhoods and interview “nicer” people. Because of this, they biased the sample in the direction of their own (largely) middle-class preferences and, as a result, under-represented constituencies that would give Truman his unexpected victory.

### IS THE SAMPLING PROCEDURE BIASED?

Because professionals using modern techniques can make bad statistical generalizations through biased sampling, it is not surprising that our everyday, informal inductive generalizations are often inaccurate. Sometimes we go astray because of small samples and biased samples. This happens, for example, when we form opinions about what people think or what people are like by asking only our friends. But bias can affect our reasoning in other ways as well.

One of the main sources of bias in everyday life is *prejudice*. Even if we sample a wide enough range of cases, we often reinterpret what we hear or see in light of some preconception. People who are prejudiced will find very little good and a great deal bad in those they despise, no matter how these people actually behave. In fact, most people are a mixture of good and bad qualities. By ignoring the former and dwelling on the latter, it is easy enough for a prejudiced person to confirm negative opinions.

Another common source of bias in sampling arises from phrasing questions in ways that encourage certain answers while discouraging others. Even if a fair sample is asked a question, it is well known that the way a question is phrased can exert a significant influence on how people will answer it. Questions like the following are not intended to elicit information, but instead to push people’s answers in one direction rather than another:

Which do you favor: (a) preserving a citizen's constitutional right to bear arms or (b) leaving honest citizens defenseless against armed criminals?

Which do you favor: (a) restricting the sale of assault weapons or (b) knuckling under to the demands of the well-financed gun lobby?

In both cases, one alternative is made to sound attractive, the other unattractive. When questions of this sort are used, it is not surprising that different pollsters can come up with wildly different results.

Now we can summarize and restate our questions. Confronted with inductive generalizations, there are four questions that we should routinely ask:

1. Are the premises acceptable?
2. Is the sample too small?
3. Is the sample biased?
4. Is the sampling procedure biased?

#### EXERCISE II

By asking the preceding questions, specify what, if anything, is wrong with the following statistical generalizations:

1. This philosophy class is about logic, so most philosophy classes are probably about logic.
2. Most college students like to ski, because I asked a lot of students at several colleges in the Rocky Mountains, and most of them like to ski.
3. K-Mart asked all of their customers throughout the country whether they prefer K-Mart to Walmart, and 90 percent said they did, so 90 percent of all shoppers in the country prefer K-Mart.
4. A Swede stole my bicycle, so most Swedes are thieves.
5. I've never tried it before, but I just put a kiwi fruit in a tub of water. It floated. So most kiwi fruits float in water.
6. I have lots of friends. Most of them think that I would make a great president. So most Americans would probably agree.
7. In exit polls after people had just voted, most people told our candidate that they voted for her, so probably most people did vote for her.
8. Mary told me that all of her older children are geniuses, so her baby will probably be a genius, too.
9. When asked whether they would prefer a tax break or a bloated budget, almost everyone said that they wanted a tax break. So a tax break is overwhelmingly popular with the people.
10. When hundreds of convicted murderers in states without the death penalty were asked whether they would have committed the murder if the state had a death penalty, most of them said that they would not have done it. So most murders can be deterred by the death penalty.

---

**STATISTICAL APPLICATIONS**

---

In a statistical generalization, we draw inferences concerning a population from information concerning a sample of that population. If 60 percent of the population sampled said that they would vote for candidate X, we might draw the conclusion that roughly 60 percent of the population will vote for candidate X. With a *statistical application* (sometimes called a *statistical syllogism*), we reason in the reverse direction: From information concerning a population, we draw a conclusion concerning a member or subset of that population. Here is an example:

Ninety-seven percent of the Republicans in California voted for Romney.  
Marvin is a Republican from California.

. . . Marvin voted for Romney.

Such arguments have the following general form:

X percent of *Fs* have the feature *G*.

*a* is an *F*.

. . . *a* has the feature *G*.<sup>1</sup>

Obviously, when we evaluate the strength of a statistical application, the percentage of *Fs* that have the feature *G* will be important. As the figure approaches 100 percent, the argument gains strength. Thus, our original argument concerning Marvin is quite strong. We can also get strong statistical applications when the figure approaches 0 percent. The following is a strong inductive argument:

Three percent of the socialists from California voted for Romney.  
Maureen is a socialist from California.

. . . Maureen did *not* vote for Romney.

Statistical applications of the kind considered here are strong only if the figures are close to 100 percent or 0 percent. When the percentages are in the middle of this range, such statistical applications are weak.

A more interesting problem in evaluating the strength of a statistical application concerns the *relevance* of the premises to the conclusion. In the above schematic representation, *F* stands for what is called the *reference class*. In our first example, being a Republican from California is the reference class; in our second example, being a socialist from California is the reference class. A striking feature of statistical applications is that using different reference classes can yield incompatible results. To see this, consider the following example:

Three percent of Obama's relatives voted for Romney.  
Marvin is a relative of Obama.

. . . Marvin did not vote for Romney.

We now have a statistical application that gives us strong support for the claim that Marvin did not vote for Romney. This is incompatible with our first statistical application, which gave strong support to the claim that he did. To overlook this conflict between arguments based on different reference classes would be a kind of fallacy. Which statistical application, if either, should we trust? This will depend on which of the reference classes we take to be more relevant. Which counts more, political affiliation or family ties? That might be hard to say.

One way of dealing with competing statistical applications is to combine the reference classes. We could ask, for example, what percentage of Republicans from California who are relatives of Obama voted for Romney? The result might come out this way:

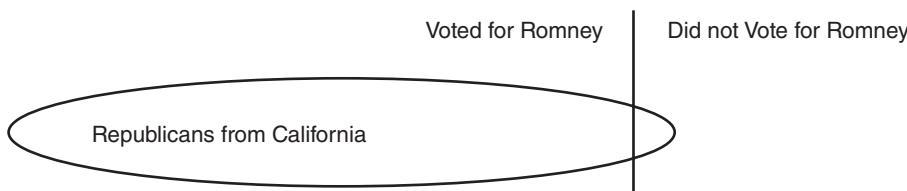
Forty-two percent of Republicans from California who were relatives of Obama voted for Romney.

Marvin is a Republican from California who is a relative of Obama.

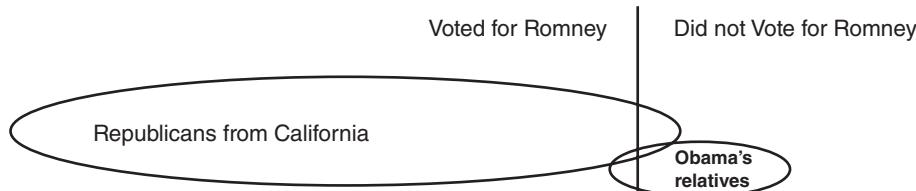
∴ Marvin voted for Romney.

This statistical application provides very weak support for its conclusion. Indeed, it supplies some weak support for the denial of its conclusion—that is, for the claim that Marvin did *not* vote for Romney.

This situation can be diagrammed with ellipses of varying sizes to represent the percentages of Californians and relatives of Obama who do or do not vote for Romney. First, we draw an ellipse to represent Republicans from California and place a vertical line so that it cuts off roughly (very roughly!) 97 percent of the area of that ellipse to represent the premise that 97 percent of the Republicans from California voted for Romney:



Next, we add a second ellipse to represent Obama's relatives:



Only about 3 percent of the small ellipse is left of the line to represent the premise that 3 percent of Obama's relatives voted for Romney. The area that lies within both ellipses represents the people who are both Republicans from California and also relatives of Obama. About 42 percent of that area is left of the line to represent the premise that 42 percent of Republicans from California who were relatives of Obama voted for Romney. The whole diagram now shows how all of these premises can be true, even though they lead to conflicting conclusions.

This series of arguments illustrates in a clear way what we earlier called the *defeasibility of inductive inferences*: A strong inductive argument can be made weak by adding further information to the premises. Given that Marvin is a Republican from California, we seemed to have good reason to think that he voted for Romney. But when we added to this the additional piece of information that he was a relative of Obama, the original argument lost most of its force. And new information could produce another reversal. Suppose we discover that Marvin, though a relative of Obama, actively campaigned for Romney. Just about everyone who actively campaigns for a candidate votes for that candidate, so it seems that we again have good reason for thinking that Marvin voted for Romney.

It is clear, then, that the way we select our reference classes will affect the strength of a statistical application. The general idea is that we should define our reference classes in a way that brings all relevant evidence to bear on the subject. But this raises difficulties. It is not always obvious which factors are relevant and which are not. In our example, party affiliation is relevant to how people voted in the 2012 election; shoe size presumably is not. Whether gender is significant, and, if so, how significant, is a matter for further statistical research.

These difficulties concerning the proper way to fix reference classes reflect a feature of all inductive reasoning: To be successful, such reasoning must take place within a broader framework that helps determine which features are significant and which features are not. Without this framework, there would be no reason not to consider shoe size when trying to decide how someone will vote. This shows how statistical applications, like all of the other inductive arguments that we will study, cannot work properly without appropriate background assumptions.

### EXERCISE III

Carry the story of Marvin two steps further, producing two more reversals in the strength of the statistical application with the conclusion that Marvin voted for Romney.

**EXERCISE IV**

For each of the following statistical applications, identify the reference class, and then evaluate the strength of the argument in terms of the percentages or proportions cited and the relevance of the reference class.

1. Less than 1 percent of the people in the world voted for Romney.  
Michelle is a person in the world.  
∴ Michelle did not vote for Romney.
2. Very few teams repeat as Super Bowl champions.  
New England was the last Super Bowl champion.  
∴ New England will not repeat as Super Bowl champion.
3. A very high percentage of people in the Senate are men.  
Elizabeth Warren is in the Senate.  
∴ Elizabeth Warren is a man.
4. Three percent of socialists with blue eyes voted for Romney.  
Maureen is a socialist with blue eyes.  
∴ Maureen did not vote for Romney.
5. Ninety-eight percent of what John says is true.  
John said that his father is also named John.  
∴ John's father is named John.
6. Ninety-eight percent of what John says is true.  
John said that the Giants are going to win.  
∴ The Giants are going to win.
7. Half the time he doesn't know what he is doing.  
He is eating lunch.  
∴ He does not know that he is eating lunch.
8. Most people do not understand quantum mechanics.  
My physics professor is a person.  
∴ My physics professor probably does not understand quantum mechanics.
9. Almost all birds can fly.  
This penguin is a bird.  
∴ This penguin can fly.
10. Most people who claim to be psychic are frauds.  
Mary claims to be psychic.  
∴ Mary is a fraud.

**DISCUSSION QUESTION**

Although both in science and in daily life, we rely heavily on the methods of inductive reasoning, this kind of reasoning raises a number of perplexing problems. The most famous problem concerning the legitimacy of induction was formulated by the eighteenth-century philosopher David Hume, first in his *Treatise of Human Nature* and then later in his *Enquiry Concerning Human Understanding*. A simplified version of Hume's skeptical argument goes as follows: Our inductive generalizations seem to rest on the assumption that *unobserved* cases will follow the patterns that we discovered in *observed* cases. That is, our inductive generalizations seem to presuppose that nature operates uniformly: The way things are observed to behave here and now are accurate indicators of how things behave anywhere and at any time. But by what right can we assume that nature is uniform? Because this claim itself asserts a contingent matter of fact, it could only be established by inductive reasoning. But because all inductive reasoning presupposes the principle that nature is uniform, any inductive justification of this principle would seem to be circular. It seems, then, that we have no ultimate justification for our inductive reasoning at all. Is this a good argument or a bad one? Why?

**NOTE**

<sup>1</sup> We can also have a *probabilistic* version of the statistical syllogism:

Ninety-seven percent of the Republicans from California voted for Romney.  
Marvin is a Republican from California.

. . . There is a 97 percent chance that Marvin voted for Romney.

We will discuss arguments concerning probability in Chapter 11.



## INFERENCE TO THE BEST EXPLANATION AND FROM ANALOGY

*Even if we can generalize from a sample to the conclusion that most women in a country voted for a certain candidate, and even if we can apply that generalization and conclude that Ilina probably voted for that candidate, this generalization and application still might not explain why Ilina voted for that candidate. Did she like his experience or his policies? Which policies? Or did she just dislike his opponent? As we saw in Chapter 1, generalization is not always enough for explanation. We also saw in Chapter 1 that some arguments can be used to explain a phenomenon when they help us understand why it happened. In contrast, explanations can also play a different role in a new kind of inductive argument. Sometimes we cite the explanatory value of a hypothesis as evidence for that hypothesis. This form of argument, which is described as inference to the best explanation, is the first topic in this chapter. It requires us to determine which explanation is best, so we will investigate common standards for assessing explanations, including falsifiability, conservativeness, modesty, simplicity, power, and depth. After explaining these standards, this chapter will turn to a related form of argument called argument from analogy, in which the fact that two things have certain features in common is taken as evidence that they have further features in common. The chapter ends by suggesting that many arguments from analogy are ultimately based on implicit inferences to the best explanation.*

### INFERENCES TO THE BEST EXPLANATION

---

One of the most common forms of inductive argument is *inference to the best explanation*.<sup>1</sup> The general idea behind such inferences is that a hypothesis gains inductive support if, when added to our stock of previously accepted beliefs, it enables us to explain something that we observe or believe, and no competing explanation works nearly as well.

To see how inferences to the best explanation work, suppose you return to your home and discover that the lock on your front door is broken and some valuables are missing. In all likelihood, you will immediately conclude

that you have been burglarized. Of course, other things *could* have produced the mess. Perhaps the police mistakenly busted into your house looking for drugs and took your valuables as evidence. Perhaps your friends are playing a strange joke on you. Perhaps a meteorite struck the door and then vaporized your valuables. In fact, all of these things *could* have happened (even the last), and further investigation could show that one of them did. Why, then, do we so quickly accept the burglary hypothesis without even considering these competing possibilities? The reason is that the hypothesis that your home was robbed is not highly improbable; and this hypothesis, together with other things we believe, provides the best—the strongest and the most natural—explanation of the phenomenon. The possibility that a meteorite struck your door is so wildly remote that it is not worth taking seriously. The possibility that your house was raided by mistake or that your friends are playing a strange practical joke on you is not wildly remote, but neither fits the overall facts very well. If it was a police raid, then you would expect to find a police officer there or at least a note. If it is a joke, then it is hard to see the point of it. By contrast, burglaries are not very unusual, and that hypothesis fits the facts extremely well. Logically, the situation looks like this:

- (1) **OBSERVATION:** Your lock is broken, and your valuables are missing.
- (2) **EXPLANATION:** The hypothesis that your house has been burglarized, combined with previously accepted facts and principles, provides a suitably strong explanation of observation 1.
- (3) **COMPARISON:** No other hypothesis provides an explanation nearly as good as that in 2.
- ∴ (4) **CONCLUSION:** Your house was burglarized.

The explanatory power of the conclusion gives us reason to believe it because doing so increases our ability to understand our observations and to make reliable predictions. Explanation is important because it makes sense out of things—makes them more intelligible—and we want to understand the world around us. Prediction is important because it tests our theories with new data and sometimes allows us to anticipate or even control future events. Inference to the best explanation enables us to achieve such goals.

Here it might help to compare inferences to the best explanation with other forms of argument. Prior to any belief about burglars, you were already *justified* in believing that your lock was broken and your valuables were missing. You could see that much. What you could not see was *why* your lock was broken. That question is what the explanation answers. Explanations help us understand why things happen, when we are already justified in believing those things did happen. (Recall Chapter 1.)

Explanations often take the form of arguments. In our example, we could argue:

- (1) Your house was burglarized.
- (2) When houses are burglarized, valuables are missing.
- ∴ (3) Your valuables are missing.

This explanatory argument starts with the hypothesis that was the conclusion of the inference to the best explanation, and it ends with the observation that was the first premise in that inference to the best explanation. The difference is that this new argument *explains* why its conclusion is true—why the valuables are missing—whereas the inference to the best explanation *justified* belief in its conclusion that your house was burglarized.

More generally, in an explanatory use of argument, we try to make sense of something by deriving it (sometimes deductively) from premises that are themselves well established. With an inference to the best explanation, we reason in the opposite direction: Instead of deriving an observation from its explanation, we derive the explanation from the observation. That a hypothesis provides the best explanation of something whose truth is already known provides evidence for the truth of that hypothesis.

Once we grasp the notion of an inference to the best explanation, we can see this pattern of reasoning everywhere. If you see your friend kick the wall, you infer that he must be angry, because there is no other explanation of why he would kick the wall. Then if he turns away when you say, “Hello,” you might think that he is angry *at you*, if you cannot imagine any other reason why he would not respond. Similarly, when your car goes dead right after a checkup, you may conclude that it is out of fuel, if that is the best explanation of why your car stopped. Psychologists infer that people care what others think about them, even when they deny it, because that explains why people behave differently in front of others than when they are alone. Linguists argue that the original Indo-European language arose millennia ago in an area that was not next to the sea but did have lakes and rivers, because that is the best explanation of why Indo-European languages have no common word for seas but do share a common root “nav-” that connotes boats or ships. Astronomers believe that our Universe began with a Big Bang, because that hypothesis best explains the background microwave radiation and spreading of galaxies. All of these arguments and many more are basically inferences to the best explanation.

Solutions to murder mysteries almost always have the form of an inference to the best explanation. The facts of the case are laid out and then the clever detective argues that, given these facts, only one person could possibly have committed the crime. In the story “Silver Blaze,” Sherlock Holmes concludes that the trainer must have been the dastardly fellow who stole Silver Blaze, the horse favored to win the Wessex Cup, which was to be run the following day. Holmes’s reasoning, as usual, was very complex, but the key part of his argument was that the dog kept in the stable did not bark loudly when someone came and took away the horse.

I had grasped the significance of the silence of the dog, for one true inference invariably suggests others. [I knew that] a dog was kept in the stables, and yet, though someone had been in and fetched out a horse, he had not barked enough to arouse the two lads in the loft. Obviously the midnight visitor was someone whom the dog knew well.<sup>2</sup>

Together with other facts, this was enough to identify the trainer, Straker, as the person who stole Silver Blaze. In this case, it is the fact that something *didn't* occur that provides the basis for an inference to the best explanation.

Of course, Holmes's inference is not absolutely airtight. It is possible that Straker is innocent and Martians with hypnotic powers over dogs committed the crime. But that only goes to show that this inference is neither valid nor deductive in our sense. It does not show anything wrong with Holmes's inference. Since his inference is inductive, it is enough for it to be strong.

Inferences to the best explanation are also defeasible. No matter how strong such an inference might be, it can always be overturned by future experience. Holmes might later find traces of a sedative in the dog's blood or someone else might confess or provide Straker with an alibi. Alternatively, Holmes (or you) might think up some better explanation. Still, unless and until such new evidence or hypothesis comes along, we have adequate reason to believe that Straker stole the horse, because that hypothesis provides the best available explanation of the information that we have now. The fact that future evidence or hypotheses always might defeat inferences to the best explanation does not show that such inferences are all bad. If it did show this, then science and everyday life would be in trouble, because so much of science and our commonsense view of the world depends on inferences to the best explanation.

## WHICH EXPLANATION IS BEST?

To assess such inferences, we still need some standards for determining which explanation is the *best*. There is, unfortunately, no simple rule for deciding this, but we can list some factors that go into the evaluation of an explanation.<sup>3</sup>

First, the hypothesis should really *explain the observations*. A good explanation makes sense out of that which it is intended to explain. In our original example, the broken lock can be explained by a burglary but not by the hypothesis that a friend came to see you (unless you have strange friends). Moreover, the hypothesis needs to explain *all* of the relevant observations. The hypothesis of a mistaken police raid might explain the broken lock but not the missing valuables or the lack of any note or police officers when you return home.

The explanation should also be *deep*. An explanation is not deep but shallow when the explanation itself needs to be explained. It does not help to explain something that is obscure by citing something just as obscure. Why did the police raid your house? Because they suspected you. That explanation is shallow if it immediately leads to another question: Why did they suspect you? Because they had the wrong address. If they did not have the wrong address, then we would wonder why they suspected you. Without an explanation of their suspicions, the police raid hypothesis could not adequately explain even the broken lock.

Third, the explanation should be *powerful*. It is a mark of excellence in an explanation that the same kind of explanation can be used successfully over a wide range of cases. Many broken locks can be explained by burglaries.

Explanatory range is especially important in science. One of the main reasons why Einstein's theory of relativity replaced Newtonian physics is that Einstein could explain a wider range of phenomena, including very small particles at very high speeds.

Explanations go too far, however, when they could explain any possible event. Consider the hypothesis that each particle of matter has its own individual spirit that makes it do exactly what it does. This hypothesis might seem to explain some phenomena that even Einstein's theory cannot explain. But the spirit hypothesis really explains nothing, because it does not explain why any particle behaves one way as opposed to another. Either behavior is compatible with the hypothesis, so neither is explained. To succeed, therefore, explanations need to be incompatible with some possible outcome. In short, they need to be *falsifiable*. (See Chapter 16 on self-sealers.)

Moreover, explanations should be *modest* in the sense that they should not claim too much—indeed, any more than is needed to explain the observations. When you find your lock broken and valuables gone, you should not jump to the conclusion that there is a conspiracy against you or that gangs have taken over your neighborhood. Without further information, there is no need to specify that there was more than one burglar in order to explain what you see. There is also no need to hypothesize that there was only one burglar. For this reason, the most modest explanation would not specify any number of burglars, so no inference to the best explanation could justify any claim about the number of burglars, at least until more evidence comes along.

Modesty is related to *simplicity*. One kind of simplicity is captured by the celebrated principle known as *Occam's razor*, which tells us not to multiply entities beyond necessity. Physicists, for example, should not postulate new kinds of subatomic particles or forces unless there is no other way to explain their experimental results. Similar standards apply in everyday life. We should not believe in ghosts unless they really are necessary to explain the noises in our attic or some other phenomenon. Simplicity is not always a matter of new kinds of entities. In comparison with earlier views, the theory that gases are composed of particles too small to see was simpler insofar as the particle theory allowed gas laws to be explained by the standard physical principles governing the motions of larger particles without having to add any new laws. Simplicity is a mark of excellence in an explanation partly because simple explanations are easier to understand and apply, but considerations of plausibility and aesthetics are also at work in judgments of which explanation is simplest.

The tests of modesty and simplicity might seem to be in tension with the test of power. This tension can be resolved only by finding the right balance. The best explanation will not claim any more than is necessary (so it will be modest), but it will claim enough to cover a wide range of phenomena (so it will be powerful). This is tricky, but the best explanations succeed in reconciling and incorporating these conflicting virtues as much as possible.

Finally, an explanation should be *conservative*. Explanations are better when they force us to give up fewer well-established beliefs. We have

strong reasons to believe that cats cannot break metal locks. This rules out the hypothesis that your neighbor's cat broke your front-door lock. Explanations should also not contain claims that are themselves too unlikely to be true. A meteorite would be strong enough to break your lock, but it is very unlikely that a meteorite struck your lock. That makes the burglary hypothesis better, at least until we find other evidence (such as meteorite fragments) that cannot be explained except by a meteorite.

In sum, a hypothesis provides the best explanation when it is more explanatory, broad, powerful, falsifiable, modest, simple, and conservative than any competing hypothesis. Each of these standards can be met to varying degrees, and they can conflict. As we saw, the desire for simplicity might have to be sacrificed to gain a more powerful explanation. Conservatism also might have to give way to explain some unexpected observations, and so on. These standards are not always easy to apply, but they can often be used to determine which explanation is best.

Once we determine that one explanation is the best, we still cannot yet infer that it is true. It might turn out that the best explanation out of a group of weak explanations isn't good enough. For centuries people were baffled by the floods that occurred in the Nile river each spring. The Nile, as far as anyone knew, flowed from an endless desert. Where, then, did the flood waters come from? Various wild explanations were suggested—mostly about deities of one kind or another—but none was any good. Looking for the best explanation among these weak explanations would be a waste of time. It was only after it was discovered that central Africa contains a high mountain range covered with snow in the winter that a reasonable explanation became possible. That, in fact, settled the matter. So it must be understood that the best explanation must also be a *good enough* explanation.

Even when an explanation is both good and best, what it explains might be illusory. Many people believe that shark cartilage prevents cancer, because the best explanation of why sharks do not get cancer lies in their cartilage. One serious problem for this inference is that sharks *do* get cancer. They even get cancer in their cartilage. So this inference to the best explanation fails.

When a particular explanation is both good and much better than any competitor, and when the explained observation is accurate, then an inference to the best explanation will provide *strong* inductive support. At other times, no clear winner or even reasonable contender emerges. In such cases, an inference to the best explanation will be correspondingly *weak*.

## CONTEXT IS CRUCIAL

Whether an inference to the best explanation is strong *enough* depends on the context. As contexts shift, standards of rigor can change. Evidence that is strong enough to justify my belief that my spouse took our car might not be strong enough to convict our neighbor of stealing our car. Good judgment is often required to determine whether a certain degree of strength is adequate for the purposes at hand.

Context can also affect the rankings of various factors. Many explanations, for example, depend on universal premises. In such cases, compatibility with observation is usually the primary test. The universal principle should not be refuted by counterexamples (see Chapter 17). But sometimes explanatory power will take precedence: If a principle has strong explanatory power, we may accept it even in the face of clear disconfirming evidence. We do not give up good explanations lightly—nor should we. One reason is that we do not test single propositions in isolation from other propositions in our system of beliefs. When faced with counterevidence to our beliefs, we often have a choice between what to give up and what to continue to hold on to. A simple example will illustrate this. Suppose that we believe the following things:

- (1) Either John or Joan committed the crime.
- (2) Whoever committed the crime must have had a motive for doing so.
- (3) Joan had no motive to commit the crime.

From these three premises we can validly infer that John committed the crime. Suppose, however, that we discover that John could not have committed the crime. (Three bishops and two judges swear that John was somewhere else at the time.) Now, from the fact that John did not commit the crime, we could not immediately conclude that Joan committed it, for that would lead to an inconsistency. If she committed the crime, then, according to premise 3, she would have committed a motiveless crime, but that conflicts with premise 2, which says that motiveless crimes do not occur. So the discovery that John did not commit the crime entails that at least one of the premises in the argument must be abandoned, but it does not tell us which one or which ones.

This same phenomenon occurs when we are dealing with counterevidence to a complex system of beliefs. Counterevidence shows that there must be something wrong somewhere in the system, but it does not show exactly where the problem lies. One possibility is that the *supposed* counterevidence is itself in error. Imagine that a student carries out an experiment and gets the result that one of the fundamental laws of physics is false. This will not shake the scientific community even a little, for the best explanation of the student's result is that she messed things up. Given well-established principles, she could not have gotten the result she did if she had run the experiment correctly. Of course, if a great many reputable scientists find difficulties with a supposed law, then the situation is different. The hypothesis that all of these scientists, like the student, simply messed up is itself highly unlikely. But it is surprising how much contrary evidence will be tolerated when dealing with a strong explanatory theory. Scientists often continue to employ a theory in the face of counterevidence. Sometimes this perpetuates errors. For years, instruments reported that the levels of ozone above Antarctica were lower than before, but scientists attributed these measurements to bad equipment, until finally they announced an ozone hole there. Still, there is often good reason to hold on to a useful theory despite counterevidence, as long as its defects do not make serious trouble—that is, give bad results in areas that count. Good judgment is required to determine when it is finally time to shift to a different explanation.

**EXERCISE I**

Imagine that you offer an explanation, and a critic responds in the following way. Which virtue (explanatoriness, depth, power, falsifiability, modesty, simplicity, or conservativeness) is your critic claiming that your explanation lacks?

1. But that won't explain anything other than this particular case.
2. But that conflicts with everything we know about biology.
3. But you don't have to claim all of that in order to explain what we see.
4. But that just raises new questions that you need to answer.
5. But that explains only a small part of the story.
6. But that would apply whatever happened.

**EXERCISE II**

For each of the following explanations, specify which standard of a good explanation, if any, it violates. The standards require that a good explanation be explanatory, deep, powerful, falsifiable, modest, simple, and conservative. A single explanation might violate more than one standard.

1. Although we usually have class at this time in this room, I don't see anybody in the classroom, because a wicked witch made them all invisible.
2. Although we usually have class at this time in this room, I don't see anybody in the classroom, because they all decided to skip class today.
3. Although we usually have class at this time in this room, I don't see anybody in the classroom, because it's Columbus Day.
4. My house fell down, because it was painted red.
5. My house fell down, because of a powerful earthquake centered on my property that did not affect anything or anybody else.
6. My house fell down, because its boards were struck by a new kind of subatomic particle.
7. Although I fished here all day, I didn't catch any fish, because there are no fish in this whole river.
8. Although I fished here all day, I didn't catch any fish, because the river gods don't like me.
9. Although I fished here all day, I didn't catch any fish, because I was unlucky today.
10. That light far up in the night sky is moving quickly, because it is the daily United Airlines flight from Boston to Los Angeles.
11. That light far up in the night sky is moving quickly, because it is an alien space ship.
12. That light far up in the night sky looks like it is moving quickly, because there's something wrong with my eyes right now.

**EXERCISE III**

Give two competing hypotheses that might be offered to explain each of the following phenomena. Which of these hypotheses is better? Why?

1. You follow a recipe carefully, but the bread never rises.
2. Your house begins to shake so violently that pictures fall off your walls.
3. Your key will not open the door of your house.
4. People start putting television cameras on your lawn, and a man with a big smile comes walking up your driveway.
5. Virtually all of the food in markets has suddenly sold out.
6. You put on a shirt and notice that there is no pocket on the front like there used to be.
7. A cave is found containing the bones of both prehistoric humans and now-extinct predators.
8. A cave is found containing the bones of both prehistoric humans and now-extinct herbivores.
9. After being visited by lobbyists for cigarette producers, your senator votes in favor of tobacco price supports, although he opposed them before.
10. Large, mysterious patterns of flattened wheat appear in the fields of Britain. (Some people attribute these patterns to visitors from another planet.)
11. A palm reader foretells that something wonderful will happen to you soon, and it does.
12. A neighbor sprinkles purple powder on his lawn to keep away tigers, and, sure enough, no tigers show up on his lawn.

**DISCUSSION QUESTIONS**

1. Read a murder mystery or detective story. Is the solution based on an inference to the best explanation? If so, put that inference in standard form, and evaluate it using the tests discussed above.
2. In the Discussion Question at the end of Chapter 1, Colin Powell gives several arguments that in 2003 Saddam Hussein was still trying to obtain fissile material for a nuclear weapons program. Which of Powell's arguments is an inference to the best explanation? How well do these arguments meet the standards for this form of argument?
3. Put the following inference to the best explanation in standard form, and then evaluate it as carefully as you can, using the tests discussed above.

*(continued)*

[During the Archean Era, which extended from about 3.8 to 2.5 million years before the present.] the sun's luminosity was perhaps 25% less than that of today. . . . This faint young sun has led to a paradox. There is no evidence from the scant rock record of the Archean that the planetary surface was frozen. However, if Earth had no atmosphere or an atmosphere of composition like that of today, the amount of radiant energy received by Earth from the sun would not be enough to keep it from freezing. The way out of this dilemma is to have an atmosphere present during the early Archean that was different in composition than that of today. . . . For a variety of reasons, it has been concluded, although still debated, that the most likely gases present in greater abundance in the Archean atmosphere were carbon dioxide, water vapor (the most important greenhouse gas) and perhaps methane. The presence of these greenhouse gases warmed the atmosphere and planetary surface and prevented the early Archean Earth from being frozen.<sup>4</sup>

4. Find three more inferences to the best explanation in articles about science in a newspaper, magazine, or professional journal. This should be easy because scientists often use this form of argument. Put those inferences in standard form, and then evaluate them using the tests discussed above.

## ARGUMENTS FROM ANALOGY

Another very common kind of inductive argument moves from a premise that two things are similar in some respects to a conclusion that they must also be analogous in a further respect. Such arguments from analogy can be found in many areas of everyday life. When we buy a new car, how can we tell whether it is going to be reliable? *Consumer Reports* might help if it is an old model; but if it is a brand-new model with no track record, then all we can go on is its similarities to earlier models. Our reasoning then seems to be that the new model is like the old model in various ways, and the old model was reliable, so the new model is probably reliable, too.

The same form of argument is used in science. Here's an example from geology:

Meteorites composed predominantly of iron provide evidence that parts of other bodies in the solar system, presumably similar in origin to Earth, were composed of metallic iron. The evidence from meteorite compositions and origins lends support to the conclusion that Earth's core is metallic iron.<sup>5</sup>

The argument here is that Earth is analogous to certain meteors in their origins, and those meteors have a large percentage of iron, so the Earth as a whole probably contains about the same percentage of iron. Because a smaller amount of iron is present in the Earth's crust, the rest must lie in the Earth's core.

Similarly, archaeologists might argue that a certain knife was used in ritual sacrifices because it resembles other sacrificial knives in its size, shape, materials, carvings, and so on. The analogy in this case is between the newly

discovered knife and the other knives. This analogy is supposed to support a conclusion about the function of the newly discovered knife.

Although such arguments from analogy have diverse contents, they share a common form that can be represented like this:

- (1) Object *A* has properties *P, Q, R*, and so on.
  - (2) Objects *B, C, D*, and so on also have properties *P, Q, R*, and so on.
  - (3) Objects *B, C, D*, and so on have property *X*.
- ∴ (4) Object *A* probably also has property *X*.

In the archaeological example, object *A* is the newly discovered knife, and objects *B, C, D*, and so on are previously discovered knives that are known to have been used in sacrifices. Properties *P, Q, R*, and so on are the size, shape, materials, and carvings that make *A* analogous to *B, C, D*, and so on. *X* is the property of being used as a sacrificial knife. Premise 3 says that the previously discovered artifacts have this property. The conclusion, on line 4, says that the newly discovered artifact probably also has this property.

Since arguments from analogy are inductive, they normally aren't valid. It is possible that, even though this knife is analogous to other sacrificial knives, this knife was used to shave the king or just to cut bread. These arguments are also defeasible. The argument about knives obviously loses all of its strength if we find "Made in China" printed on the newly discovered knife. Still, none of this shows that arguments from analogy are no good. Despite being invalid and defeasible, some arguments from analogy can still provide reasons—even strong reasons—for their conclusions.

How can we tell whether an argument from analogy is strong or weak? One obvious requirement is that the premises must be *true*. If the previously discovered knives were not really used in sacrifices, or if they do not really have the same carvings on their handles as the newly discovered knife, then this argument from analogy does not provide much, if any, support for its conclusion.

In addition, the cited similarities must be *relevant*. Suppose someone argues that his old car was red with a black interior and had four doors and a sunroof, and his new car also has these properties, so his new car is probably going to be as reliable as his old car. This argument is very weak because the cited similarities are obviously irrelevant to reliability. Such assessments of relevance depend on background beliefs, such as that reliability depends on the drive train and the engine rather than on the color or the sunroof.

The similarities must also be *important*. Similarities are usually more important the more specific they are. Lots of cars with four tires and a motor are reliable, but this is not enough to infer that, because this particular car also has four tires and a motor, it will be reliable, too. The reason is obvious: There are also lots of unreliable cars with four tires and a motor. In general, if many objects have properties *P, Q*, and *R*, and many of those lack property *X*, then arguments from these analogies will be weak. In contrast, if a

smaller percentage of objects that have properties *P*, *Q*, and *R* lack property *X*, then the argument from these analogies will be strong.

If we are not sure which respects *are* important, we still might have some idea of which respects *might* be important. Then we can try to cite objects that are analogous in as many as possible of those respects. By increasing the number of potentially relevant respects for which the analogy holds, we can increase the likelihood that the important respects will be on our list. That shows why arguments from analogy are usually stronger when they cite *more and closer analogies* between the objects.

Another factor that affects the strength of an argument from analogy is the presence of *relevant disanalogies*. Because arguments from analogy are defeasible, as we saw, a strong argument from analogy can become weak if we add a premise that states an important disanalogy. Suppose my new car is like my old cars in many ways, but there is one difference: The new car has an electric motor, whereas the old cars were powered by gasoline. This one difference is enough to weaken any argument to the conclusion that the new car will be reliable. Of course, other disanalogies, such as a different color, won't matter to reliability; and it will often require background knowledge to determine how important a disanalogy is.

We need to be careful here. Some disanalogies that are relevant do not undermine an argument from analogy. If a new engine design was introduced by top engineers to increase reliability, then this disanalogy might not undermine the argument from analogy. Differences that point to more reliability rather than less might even make the argument from analogy stronger.

Other disanalogies can increase the strength of an argument from analogy in a different way. If the same markings are found on very different kinds of sacrificial knives, then the presence of those markings on the newly discovered knife is even stronger evidence that this knife was also used in sacrifices. Differences among the cases cited only in the premises as analogies (that is, *B*, *C*, *D*, and so on) can strengthen an argument from analogy.

Finally, the strength of an argument from analogy depends on its conclusion. Analogies to other kinds of cars provide stronger evidence for a weak conclusion (such as that the new model will probably be pretty reliable) and weaker evidence for a strong conclusion (such as that the new model will definitely be just as reliable as the old model). As with other forms of argument, an argument from analogy becomes stronger as its conclusion becomes weaker and vice versa.

These standards can be summarized by saying that an argument from analogy is stronger when:

1. It cites more and closer analogies that are more important.
2. There are fewer or less important disanalogies between the object in the conclusion and the other objects.
3. The objects cited only in the premises are more diverse.
4. The conclusion is weaker.

## ARE ANALOGIES EXPLANATIONS?

After learning about arguments from analogy, it is natural to wonder how they are related to inferences to the best explanation. Although this is sometimes disputed, it seems to us that arguments from analogy are often—if not always—implicit and incomplete inferences to the best explanation. As we pointed out, analogies don't support any conclusion unless they are relevant, and whether they are relevant depends on how they fit into explanations. The color of a car is irrelevant to its reliability, because color plays no role in explaining its reliability. What explains its reliability is its drive train design, materials, care in manufacturing, and so on. That is why analogies in those respects can support a conclusion about reliability. Similarly, the markings on an artifact are relevant to whether it is a sacrificial knife if the best explanation of why it has those markings is that it was used in sacrifices. What makes that explanation best is that it also explains similar markings on other sacrificial knives. Thus, such arguments from analogy can be seen as involving an inference to the best explanation of why objects *B*, *C*, *D*, and so on have property *X* followed by an application of that explanation to the newly discovered object *A*.

Sometimes the explanation runs in the other direction. Whereas the conclusion about the knife's use (*X*) is supposed to explain its shared markings (*P*, *Q*, *R*), sometimes it is the shared features (*P*, *Q*, *R*) that are supposed to explain the feature claimed in the conclusion (*X*). Here is a classic example:

We may observe a very great [similarity] between this earth which we inhabit, and the other planets, Saturn, Jupiter, Mars, Venus, and Mercury. They all revolve around the sun, as the earth does, although at different distances and in different periods. They borrow all their light from the sun, as the earth does. Several of them are known to revolve around their axis like the earth, and, by that means, must have a like succession of day and night. Some of them have moons that serve to give them light in the absence of the sun, as our moon does to us. They are all, in their motions, subject to the same law of gravitation, as the earth is. From all this similarity it is not unreasonable to think that those planets may, like our earth, be the habitation of various orders of living creatures. There is some probability in this conclusion from analogy.<sup>6</sup>

The argument here seems to be that some other planet probably supports life, because Earth does and other planets are similar to Earth in revolving around the sun and around an axis, getting light from the sun, and so on. What makes certain analogies relevant is not, of course, that the motion of Earth is explained by the presence of life here. Rather, certain features of Earth explain why Earth is habitable. The argument suggests that the best explanation of why there is life on our planet is that certain conditions make life possible. That generalization can then be used to support the conclusion that other planets with the same conditions probably support life as well.

In one way or another, many (or maybe even all) arguments from analogy can be seen as inferences to the best explanation. But they are usually

incomplete explanations. The argument for life on other planets did not have to commit itself to any particular theory about the origin of life or about which conditions are needed to support life. Nor did the car argument specify exactly what makes cars reliable. Such arguments from analogy merely list a number of similarities so that the list will be likely to include whatever factors are needed for life or for reliability. In this way, arguments from analogy can avoid depending on any complete theory about what is and what is not relevant.

This incompleteness makes arguments from analogy useful in situations where we do not yet know enough to formulate detailed theories or even to complete an inference to the best explanation. Yet, the incompleteness of arguments from analogy also makes them more vulnerable to refutation, since the analogies that they list might fail to include a crucial respect. This does not mean that arguments from analogy are never any good. They can be strong. However, it does suggest that their strength will increase as they approach or approximate more complete inferences to the best explanation.

#### EXERCISE IV

For each of the following arguments, state whether the indicated changes would make the argument weaker or stronger, and explain why. The strength of the argument might not be affected at all. If so, say why it is not affected.

1. My friend and I have seen many movies together, and we have always agreed on whether they are good or bad. My friend liked the movie trilogy *The Lord of the Rings*. So I probably will like it as well.

Would this argument be weaker or stronger if:

- a. The only movies that my friend and I have watched together are comedies, and *The Lord of the Rings* is not a comedy.
- b. My friend and I have seen very many, very different movies together.
- c. My friend and I always watched movies together on Wednesdays, but my friend watched *The Lord of the Rings* on a weekend.
- d. The conclusion claims that I definitely will like *The Lord of the Rings* a lot.
- e. The conclusion claims that I probably won't totally dislike *The Lord of the Rings*.

2. All the students from Joe's high school with high grades and high board scores did well in college. Joe also had high grades and board scores. So he will probably do well in college.

Would this argument be weaker or stronger if:

- a. The other students worked hard, but Joe's good grades came easily to him, so he never learned to work hard.
- b. Joe is going to a different college than the students with whom he is being compared.

- c. Joe plans to major in some easy subject, but the other students were pre-med.
  - d. Joe recently started taking drugs on a regular basis.
  - e. Joe needs to work full-time to pay his college expenses, but the others had their expenses paid by their parents.
3. A new drug cures a serious disease in rats. Rats are similar to humans in many respects. Therefore, the drug will probably cure the same disease in humans.

Would this argument be weaker or stronger if:

- a. The disease affects the liver, and rat livers are very similar to human livers.
- b. The drug does not cure this disease in cats.
- c. The drug has to be injected into the rat's tail to be effective (that is, it does not work if it is injected anywhere else in the rat).
- d. No drug of this general type has been used on humans before.
- e. The effects of the drug are enhanced by eating cooked foods.

#### EXERCISE V

Using the criteria mentioned above, evaluate each of the following arguments as strong or weak. Explain your answers. Be sure to specify the properties on which the analogy is based, as well as any background beliefs on which your evaluation depends.

1. This landscape by Cézanne is beautiful. He did another painting of a similar scene around the same time. So it is probably beautiful, too.
2. My aunt had a Siamese cat that bit me, so this Siamese cat will probably bite me, too.
3. The students I know who took this course last year got grades of A. I am a lot like them, since I am also smart and hardworking; and the course this year covers very similar material. So I will probably get an A.
4. This politician was caught cheating in his marriage, and he will have to face similarly strong temptations in his public duties, so he will probably cheat in political life as well.
5. A very high minimum wage led to increased unemployment in one country. That country's economy is similar to the economy in a different country. So a very high minimum wage will probably lead to increased unemployment in the other country as well.
6. I feel pain when someone hits me hard on the head with a baseball bat. Your body is a lot like mine. So you would probably feel pain if I hit you hard on the head with a baseball bat. (This is related to the so-called Problem of Other Minds.)

7. It is immoral for a doctor to lie to a patient about a test result, even if the doctor thinks that lying is in the patient's best interest. We know this because even doctors would agree that it would be morally wrong for a financial adviser to lie to them about a potential investment, even if the financial advisor thinks that this lie is in the doctor's best interests.
8. Chrysler was held legally liable for damages due to defects in the suspension of its Corvair. The defects in the Pinto gas tank caused injuries that were just as serious. Thus, Ford should also be held legally liable for damages due to those defects.

#### DISCUSSION QUESTIONS

1. The following excerpt presents evidence that Neanderthals were cannibals. Put the central argument from analogy, which is italicized here, into standard form. Then reconstruct the argument as an inference to the best explanation. Which representation best captures the force of the argument, or are they equally good?

#### “A GNAWING QUESTION IS ANSWERED”<sup>7</sup>

Tim White is worried that he may have helped to pin a bad rap on the Neanderthals, the prehistoric Europeans who died out 25,000 years ago. “There is a danger that everyone will think that all Neanderthals were cannibals and that’s not necessarily true,” he says. White was part of a French-American team of paleoanthropologists who recently found conclusive evidence that at least some Neanderthals ate others about 100,000 years ago. But that doesn’t mean they were cannibalistic by nature, he stresses. Most people don’t realize that cannibalism is widespread throughout nature, says White, a professor at the University of California at Berkeley and the author of a book on prehistoric cannibalism.

The question of whether the Neanderthals were cannibals had long been a hotly debated topic among anthropologists. No proof had ever been found. That debate ended, however, with the recent analysis by the team of stone tools and bones found in a cave at Moula-Guercy in southern France. The cave is about the size of a living room, perched about 80 metres above the Rhone River. “This one site has all of the evidence right together,” says White. “It’s as if somebody put a yellow tape around the cave for 100,000 years and kept the scene intact.” The bones of deer and other fauna show the clear markings of the nearby stone tools, indicating the deer had been expertly butchered; they were skinned, their body parts cut off and the meat and tendons sliced from

Source: A Gnawing Question is Answered by Michael Downey as appeared in THE TORONTO STAR, October 10, 1999.

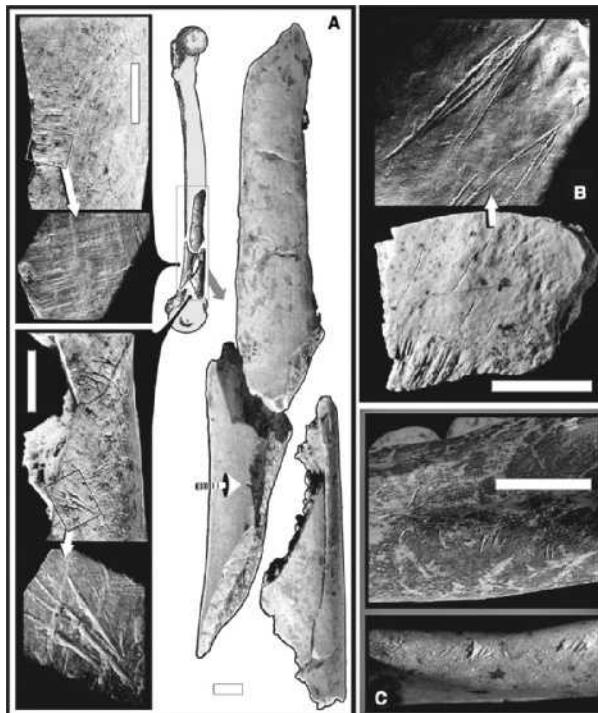
the bone. Long bones were bashed open "to get at the fatty marrow inside," says White.

So what does all this have to do with cannibalism? *The bones of the six (so far) humans in the same locations have precisely the same markings made by the same tools. That means these fairly modern humans were skinned and eaten in the same manner as the deer.*

And if you are thinking they were eaten after they just happened to die, they do represent all age groups. Two were children about 6 years old, two were teenagers, and two were adults.

But maybe they were eaten at a time when food was unusually scarce, right? Not so. There is a large number of animal bones at the same dig, indicating that there were options to eating other Neanderthals.

Human bones with similar cut marks have been found throughout Europe, from Spain to Croatia, providing tantalizing hints of Neanderthal cannibalism activity over tens of thousands of years. But finding such clear evidence of the same preparation techniques being used on deer in the same cave site in France, will "necessitate reassessment of earlier finds," always attributed to ritual burial practices or some other explanation, says White.



From DEFLEUR ET AL., SCIENCE 286:128 (1999). Reprinted with permission from AAAS.

(continued)

2. In the following passage, William Paley argues for the existence of God on the basis of an analogy to a watch. Reconstruct this argument from analogy and then evaluate it by applying the criteria discussed above. Could Paley's argument also be reconstructed as an inference to the best explanation? If so, would that reconstruction better capture the force of the argument?

### **“THE WATCH AND THE WATCHMAKER”<sup>8</sup>**

In crossing a heath, suppose I pitched my foot against a stone and were asked how the stone came to be there, I might possibly answer that for anything I knew to the contrary it had lain there forever; nor would it, perhaps, be very easy to show the absurdity of this answer. But suppose I had found a watch upon the ground, and it should be inquired how the watch happened to be in that place, I should hardly think of the answer which I had before given, that for anything I knew the watch might have always been there. Yet why should not this answer serve for the watch as well as for the stone? Why is it not as admissible in the second case as in the first? For this reason, and for no other, namely, that when we come to inspect the watch, we perceive—what we could not discover in the stone—that its several parts are framed and put together for a purpose, e.g., that they are so formed and adjusted as to produce motion, and that motion so regulated as to point out the hour of the day; that if the different parts had been differently shaped from what they are, of a different size from what they are, or placed after any other manner or in any other order than that in which they are placed, either no motion at all would have been carried on in the machine, or none which would have answered the use that is now served by it. . . . This mechanism being observed—it requires indeed an examination of the instrument, and perhaps some previous knowledge of the subject, to perceive and understand it; but being once, as we have said, observed and understood—the inference we think is inevitable, that the watch must have had a maker—that there must have existed, at some time and at some place or other, an artificer or artificers who formed it for the purpose which we find it actually to answer, who comprehended its construction and designed its use. . . .

[E]very indication of contrivance, every manifestation of design, which existed in the watch, exists in the works of nature; with the difference, on the side of nature, of being greater and more, and that in a degree which exceeds all computation. I mean that the contrivances of nature surpass the contrivances of art, in the complexity, subtlety, and curiosity of the mechanism; and still more, if possible, do they go beyond them in number and variety; yet in a multitude of cases, are not less evidently mechanical, not less evidently contrivances, not less evidently accommodated to their end, or suited to their office, than are the most perfect productions of human ingenuity.

I know no better method of introducing so large a subject, than that of comparing a single thing with a single thing: an eye, for example, with a telescope. As far as the examination of the instrument goes, there is precisely

the same proof that the eye was made for vision, as there is that the telescope was made for assisting it. They are made upon the same principles; both being adjusted to the laws by which the transmission and refraction of rays of light are regulated. I speak not of the origin of the laws themselves; but such laws being fixed, the construction in both cases is adapted to them. . . .

To some it may appear a difference sufficient to destroy all similitude between the eye and the telescope, that the one is a perceiving organ, the other an unperceiving instrument. The fact is that they are both instruments. And as to the mechanism, at least as to mechanism being employed, and even as to the kind of it, this circumstance varies not the analogy at all. . . . The end is the same; the means are the same. The purpose in both is alike; the contrivance for accomplishing that purpose is in both alike. The lenses of the telescopes, and the humors of the eye, bear a complete resemblance to one another, in their figure, their position, and in their power over the rays of light, viz. in bringing each pencil to a point at the right distance from the lens; namely, in the eye, at the exact place where the membrane is spread to receive it. How is it possible, under circumstances of such close affinity, and under the operation of equal evidence, to exclude contrivance from the one; yet to acknowledge the proof of contrivance having been employed, as the plainest and clearest of all propositions, in the other? . . .

Were there no example in the world of contrivance except that of the *eye*, it would be alone sufficient to support the conclusion, which we draw from it, as to the necessity of an intelligent Creator. . . . The proof is not a conclusion that lies at the end of a chain of reasoning, of which chain each instance of contrivance is only a link, and of which, if one link fail, the whole fails; but it is an argument separately supplied by every separate example. An error in stating an example affects only that example. The argument is cumulative in the fullest sense of that term. The eye proves it without the ear; the ear without the eye. The proof in each example is complete; for when the design of the part and the conduciveness of its structure to that design is shown, the mind may set itself at rest; no further consideration can detract anything from the force of the example.

## NOTES

<sup>1</sup> Gilbert Harman deserves much credit for calling attention to the importance of inferences to the best explanation; see, for example, his *Thought* (Princeton, NJ: Princeton University Press, 1973). A similar form of argument called abduction was analyzed long ago by Charles Sanders Peirce; see, for example, his *Collected Papers of Charles Sanders Peirce*, Vol. 5 (Cambridge, MA: Harvard University Press, 1931), 189. A wonderful recent discussion is Peter Lipton, *Inference to the Best Explanation* (London: Routledge, 1991).

<sup>2</sup> Sir Arthur Conan Doyle, "Silver Blaze," *The Complete Sherlock Holmes*, Vol. 1 (Garden City, NY: Doubleday, 1930), 349. The stories describe Holmes as a master of deduction, but his arguments are inductive as we define the terms.

<sup>3</sup> This discussion in many ways parallels and is indebted to the fifth chapter of W. V. Quine and J. S. Ullian, *The Web of Belief*, 2nd ed. (New York: Random House, 1978).

<sup>4</sup> From Fred T. Mackenzie, *Our Changing Planet* (Upper Saddle River, NJ: Prentice-Hall, 1998), 192.

<sup>5</sup> Ibid., 42.

<sup>6</sup> Thomas Reid, *Essays on the Intellectual Powers of Man* (Cambridge, MA: MIT Press, 1969), essay I, section 4, 48.

<sup>7</sup> From Michael Downey, *The Toronto Star*, October 10, 1999.

<sup>8</sup> From William Paley, *Natural Theology* (New York: Hopkins, 1836).

## CAUSAL REASONING

*One common way to explain a phenomenon is to cite its cause. You can understand why your clothes shrunk by learning what caused them to shrink. Was the water too hot when you washed them, did you dry them too long, or was it some combination of factors? In order to determine what causes what, we need to engage in a new kind of inductive reasoning—causal reasoning—which is the topic of this chapter. Causal reasoning is often based on negative and positive tests for necessary conditions and for sufficient conditions. After developing these tests and applying them to a concrete example, we will discuss concomitant variation as a method of drawing causal conclusions from imperfect correlations. Our goal throughout this chapter is to improve our ability to identify causes so that we can better understand why certain effects happened and also make better predictions about whether similar events will happen in the future.*

### REASONING ABOUT CAUSES

---

If our car goes dead in the middle of rush-hour traffic just after its 20,000-mile checkup, we assume that there must be some reason why this happened. Cars just don't stop for no reason at all. So we ask, "What caused our car to stop?" The answer might be that it ran out of gas. If we find, in fact, that it did run out of gas, then that will usually be the end of the matter. We will think that we have discovered why this particular car stopped running. This reasoning is about a particular car on a particular occasion, but it rests on certain *generalizations*: We are confident that our car stopped running when it ran out of gas, because we believe that all cars stop running when they run out of gas. We probably did not think about this, but our causal reasoning in this particular case appealed to a commonly accepted causal generalization: Lack of fuel causes cars to stop running. Many explanations depend on *causal generalizations*.

Causal generalizations are also used to *predict* the consequences of particular actions or events. A race car driver might wonder, for example, what would happen if he added just a bit of nitroglycerin to his fuel mixture. Would it give him better acceleration, blow him up, do very little, or what?

In fact, the driver may not be in a position to answer this question straight off, but his thinking will be guided by the causal generalization that igniting nitroglycerin can cause a dangerous explosion.

So a similar pattern arises for both causal explanation and causal prediction. These inferences contain two essential elements:

1. The facts in the particular case. (For example, the car stopped and the gas gauge reads empty; or I just put a pint of nitroglycerin in the gas tank of my Maserati, and I am about to turn the ignition key.)
2. Certain causal generalizations. (For example, cars do not run without gas, or nitroglycerin explodes when ignited.)

The basic idea is that causal inferences bring particular facts under causal generalizations.

This shows why causal generalizations are important, but what exactly are they? Although this issue remains controversial, here we will treat them as a kind of *general conditional*. A general conditional has the following form:

For all  $x$ , if  $x$  has the feature  $F$ , then  $x$  has the feature  $G$ .

We will say that, according to this conditional,  $x$ 's having the feature  $F$  is a *sufficient condition* for its having the feature  $G$ ; and  $x$ 's having the feature  $G$  is a *necessary condition* for its having the feature  $F$ .

Some general conditionals are not *causal*. Neither of these two general conditionals expresses a causal relationship:

If something is a square, then it is a rectangle.

If you are eighteen years old, then you are eligible to vote.

The first conditional tells us that being a square is sufficient for being a rectangle, but this is a *conceptual* (or *a priori*) relationship, not a causal one. The second conditional tells us that being eighteen years old is a sufficient condition for being eligible to vote. The relationship here is *legal*, not causal.

Although many general conditionals are not causal, all causal conditionals are general, in our view. Consequently, if we are able to show that a causal conditional is false just by virtue of its being a general conditional, we will have refuted it. This will serve our purposes well, for in what follows we will be largely concerned with finding reasons for *rejecting* causal generalizations.

It is important to weed out false causal generalizations, because they can create lots of trouble. Doctors used to think that bloodletting would cure disease. They killed many people in the process of trying to heal them. Thus, although we need causal generalizations for getting along in the world, we also need to get them right. We will be more likely to succeed if we have proper principles for testing and applying such generalizations.

In the past, very elaborate procedures have been developed for this purpose. The most famous set of such procedures was developed by John Stuart Mill and has come to be known as Mill's methods.<sup>1</sup> Though inspired

by Mill's methods, the procedures introduced here involve some fundamental simplifications; whereas Mill introduced five methods, we will introduce only three primary rules.

The first two rules are the sufficient condition test (SCT) and the necessary condition test (NCT). We will introduce these tests first at an abstract level. One advantage of formulating these tests abstractly is so that they can be applied to other kinds of sufficient and necessary conditions, for example, those that arise in legal and moral reasoning. Once it is clear how these tests work in general, we will apply them specifically to causal reasoning.

## SUFFICIENT CONDITIONS AND NECESSARY CONDITIONS

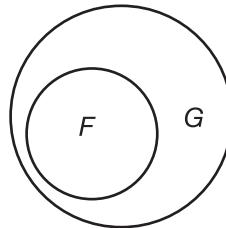
To keep our discussion as general as possible, we will adopt the following definitions of sufficient conditions and necessary conditions:

Feature *F* is a *sufficient* condition for feature *G* if and only if anything that *has* feature *F* also *has* feature *G*.

Feature *F* is a *necessary* condition for feature *G* if and only if anything that *lacks* feature *F* also *lacks* feature *G*.

These definitions are equivalent to those in the previous section, because, if anything that *lacks* feature *F* also *lacks* feature *G*, then anything that *has* feature *G* must also *have* feature *F*; and if anything that *has* feature *G* must also *have* feature *F*, then anything that *lacks* feature *F* also *lacks* feature *G*. It follows that feature *F* is a sufficient condition for feature *G* if and only if feature *G* is a necessary condition for feature *F*.

When *F* is sufficient for *G*, the relation between these features can be diagrammed like this:



The inside circle represents the sufficient condition, because anything inside that inside circle must also be inside the outside circle. The outside circle represents the necessary condition, for anything outside the outside circle must also be outside the inside circle.

These diagrams, along with the preceding definitions, should make it clear that something can be a sufficient condition for a feature without being a necessary condition for that feature, and vice versa. For example, being the element mercury is a *sufficient* condition for being a metal, but it is not a *necessary* condition for being a metal, since there are other metals. Similarly, being a metal is a *necessary* condition for being mercury, but it is not a *sufficient* condition for being mercury. Of course, some necessary conditions are also sufficient conditions. Being mercury is both necessary and sufficient for being a metallic element that is liquid at twenty degrees Centigrade. Nonetheless, many necessary conditions are not sufficient conditions, and vice versa, so we need to be careful not to confuse the two kinds of conditions.

This distinction becomes complicated when conditions get complex. Our definitions and tests hold for all features, whether positive or negative (such as not having hair) and whether simple or conjunctive (such as having both a beard and a mustache) or disjunctive (such as having either a beard or a mustache). Thus, not having any hair (anywhere) on your head is a sufficient condition of not having a beard, so not having a beard is a necessary condition of not having any hair on your head. But not having any hair on your head is not necessary for not having a beard, because you can have some hair on the top of your head without having a beard. Negation can create confusion, so we need to think carefully about what is being claimed to be necessary or sufficient for what.

Even in simple cases without negation, conjunction, or disjunction, there is a widespread tendency to confuse necessary conditions with sufficient conditions. It is important to keep these concepts straight, for, as we will see, the tests concerning them are fundamentally different.

### EXERCISE I

Which of the following claims are true? Which are false?

1. Being a car is a sufficient condition for being a vehicle.
2. Being a car is a necessary condition for being a vehicle.
3. Being a vehicle is a sufficient condition for being a car.
4. Being a vehicle is a necessary condition for being a car.
5. Being an integer is a sufficient condition for being an even number.
6. Being an integer is a necessary condition for being an even number.
7. Being an integer is a sufficient condition for being either an even number or an odd number.
8. Being an integer is a necessary condition for being either an even number or an odd number.
9. Not being an integer is a sufficient condition for not being an odd number.

10. Not being an integer is a sufficient condition for not being an even number.
11. Being both an integer and divisible by 2 without remainder is a sufficient condition for being an even number.
12. Being both an integer and divisible by 2 without remainder is a necessary condition for being an even number.
13. Being an integer divisible by 2 without remainder is a necessary condition for being an even number.
14. Driving seventy-five miles per hour (for fun) is a sufficient condition for violating a legal speed limit of sixty-five miles per hour.
15. Driving seventy-five miles per hour (for fun) is a necessary condition for violating a legal speed limit of sixty-five miles per hour.
16. Cutting off Joe's head is a sufficient condition for killing him.
17. Cutting off Joe's head is a necessary condition for killing him.
18. Cutting off Joe's head and then holding his head under water for ten minutes is a sufficient condition for killing him.

### EXERCISE II

Indicate whether the following principles are true or false and why.

1. If having feature *F* is a *sufficient* condition for having feature *G*, then having feature *G* is a *necessary* condition for having feature *F*.
2. If having feature *F* is a *sufficient* condition for having feature *G*, then lacking feature *F* is a *necessary* condition for lacking feature *G*.
3. If lacking feature *F* is a *sufficient* condition for having feature *G*, then having feature *F* is a *necessary* condition for lacking feature *G*.
4. If lacking feature *F* is a *sufficient* condition for having feature *G*, then lacking feature *F* is a *necessary* condition for having feature *G*.
5. If having either feature *F* or feature *G* is a *sufficient* condition for having feature *H*, then having feature *F* is a *sufficient* condition for having feature *H*.
6. If having either feature *F* or feature *G* is a *sufficient* condition for having feature *H*, then having feature *G* is a *sufficient* condition for having feature *H*.
7. If having either feature *F* or feature *G* is a *sufficient* condition for having feature *H*, then not having feature *F* is a *necessary* condition for not having feature *H*.
8. If having both feature *F* and feature *G* is a *necessary* condition for having feature *H*, then lacking feature *F* is a *sufficient* condition for lacking feature *H*.
9. If not having both feature *F* and feature *G* is a *sufficient* condition for having feature *H*, then lacking feature *F* is a *sufficient* condition for having feature *H*.
10. If having either feature *F* or feature *G* is a *sufficient* condition for having feature *H*, then having both feature *F* and feature *G* is a *sufficient* condition for having feature *H*.

## THE SUFFICIENT CONDITION TEST

We can now formulate tests to determine when something meets our definitions of sufficient conditions and necessary conditions. It will simplify matters if we first state these tests formally using letters. We will also begin with a simple case where we consider only four *candidates*—*A*, *B*, *C*, and *D*—for sufficient conditions for a *target* feature, *G*. *A* will indicate that the feature is present;  $\sim A$  will indicate that this feature is absent. Using these conventions, suppose that we are trying to decide whether any of the four features—*A*, *B*, *C*, or *D*—could be a sufficient condition for *G*. To this end, we collect data of the following kind:

TABLE 1

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>G</i>
Case 1:	<i>A</i>				
Case 2:	$\sim A$	<i>B</i>	<i>C</i>	$\sim D$	$\sim G$
Case 3:	<i>A</i>	$\sim B$	$\sim C$	$\sim D$	$\sim G$

We know by definition that, for something to be a sufficient condition of something else, when the former is present, the latter must be present as well. Thus, to test whether a candidate really is a sufficient condition of *G*, we only have to examine cases in which the target feature, *G*, is absent, and then check to see whether any of the candidate features are present. The sufficient condition test (SCT) can be stated as follows:

**SCT:** Any candidate that is present when *G* is absent is eliminated as a possible sufficient condition of *G*.

The test applies to Table 1 as follows: Case 1 need not be examined because *G* is present, so there can be no violation of SCT in Case 1. Case 2 eliminates two of the candidates, *B* and *C*, for both are present in a situation in which *G* is absent. Finally, Case 3 eliminates *A* for the same reason. We are thus left with *D* as our only remaining candidate for a sufficient condition for *G*.

Now let's consider feature *D*. Having survived the application of the SCT, does it follow that *D* is a sufficient condition for *G*? No! On the basis of what we have been told so far, it remains entirely possible that the discovery of a further case will reveal an instance where *D* is present and *G* absent, thus showing that *D* is also not a sufficient condition for *G*.

Case 4:       $\sim A$       *B*      *C*      *D*       $\sim G$

In this way, it is always possible for new cases to refute any inference from a limited group of cases to the conclusion that a certain candidate is a sufficient condition. In contrast, no further case can change the fact that *A*, *B*, and *C* are not sufficient conditions, because they fail the SCT.

This observation shows that, when we apply the SCT to rule out a candidate as a sufficient condition, our argument is *deductive*. We simply find a counterexample to the universal claim that a certain feature is sufficient. (See Chapter 17 on counterexamples.) However, when a candidate is not ruled out and we draw the positive conclusion that that candidate *is* a sufficient condition, then our argument is *inductive*. Inductive inferences, however well

confirmed, are always defeasible. (Recall Captain Cook's discovery of black swans at the start of Chapter 8.) That is why our inductive inference to the conclusion that  $D$  is a sufficient condition could be refuted by the new data in Case 4.

## THE NECESSARY CONDITION TEST

The necessary condition test (NCT) is like the SCT, but it works in the reverse fashion. With the SCT, we eliminated a candidate  $F$  from being the sufficient condition for  $G$ , if  $F$  was ever present when  $G$  was absent. With the NCT, we eliminate a candidate  $F$  from being a necessary condition for  $G$  if we can find a case where  $G$  is present, but  $F$  is not. This makes sense, because if  $G$  can be present when  $F$  is not, then  $F$  cannot be necessary for the occurrence of  $G$ . Thus, in applying the necessary condition test, we only have to examine cases in which the target feature,  $G$ , is present, and then check to see whether any of the candidate features are absent.

**NCT:** Any candidate that is absent when  $G$  is present is eliminated as a possible necessary condition of  $G$ .

The following table gives an example of an application of this test:

TABLE 2

Case 1:	$A$	$B$	$C$	$D$	$\sim G$
Case 2:	$\sim A$	$B$	$C$	$D$	$G$
Case 3:	$A$	$\sim B$	$C$	$\sim D$	$G$

Because Case 1 does not provide an instance where  $G$  is present, it cannot eliminate any candidate as a necessary condition of  $G$ . Case 2 eliminates  $A$  as a necessary condition of  $G$ , since it shows that  $G$  can be present without  $A$  being present. Case 3 then eliminates both  $B$  and  $D$ , leaving  $C$  as the only possible candidate for being a necessary condition for  $G$ .

From this, of course, it does not follow that  $C$  is a necessary condition for  $G$ , for, as always, new cases might eliminate it as well. The situation is the same as with the SCT. An argument for a negative conclusion that a candidate is not a necessary condition, because that candidate fails the NCT, is a deductive argument that cannot be overturned by any further cases. In contrast, an argument for a positive conclusion that a candidate is a necessary condition, because that candidate passes the NCT, is an inductive argument that can be overturned by a further case where this candidate fails the NCT. For example, suppose we find:

Case 4:     $\sim A$      $\sim B$      $\sim C$      $\sim D$      $G$

The information in this new Case 4 is enough to show that  $C$  cannot be a necessary condition of the target feature  $G$ , regardless of what we found in Cases 1–3.

In applying both the SCT and the NCT, it is crucial to specify the target feature. Case 4 shows that candidate  $C$  is not a necessary condition for target feature  $G$ . Nonetheless, candidate  $C$  still might be necessary for the

opposite target feature,  $\sim G$ . It also might be necessary for features  $A$ ,  $B$ , and  $D$ . Nothing in Cases 1–4 rules out these possibilities. Thus, even after Case 4, we cannot say simply that  $C$  is not a necessary condition. Case 4 shows that candidate feature  $C$  is not a necessary condition for target feature  $G$ , but  $C$  still might be necessary for something else. The same point applies to sufficient conditions as well. In Table 1, Case 2 ruled out the possibility that candidate feature  $B$  is sufficient for target feature  $G$ , but none of the cases in Table 1 show that  $B$  is not sufficient for target feature  $C$ . To avoid confusion, then, it is always important to specify the target feature when talking about what is or is not a necessary or sufficient condition.

### THE JOINT TEST

It is also possible to apply these rules simultaneously in the search for possible conditions that are both sufficient and necessary. Any candidate cannot be both sufficient and necessary if it fails either the SCT or the NCT. In Table 2,  $C$  is the only possible necessary condition for  $G$ , and it is not also a possible sufficient condition for  $G$ , since  $C$  fails the SCT in Case 1, where  $C$  is present and  $G$  is absent. In Table 1, however,  $D$  is a possible sufficient condition of  $G$ , because  $D$  is never present when  $G$  is absent; and  $D$  might also be a necessary condition for  $G$ , since  $G$  is never present when  $D$  is absent. Thus, none of Cases 1–3 in Table 1 eliminates  $D$  as a candidate for a condition that is both sufficient and necessary for  $G$ . As before, this possibility still might be refuted by Case 4, so any inference to a positive conclusion that some candidate is a necessary and sufficient condition must be defeasible and, hence, inductive.

### EXERCISE III

For each of the following tables determine

- Which, if any, of the candidates— $A$ ,  $B$ ,  $C$ , or  $D$ —is not eliminated by the *sufficient* condition test as a sufficient condition for target feature  $G$ ?
- Which, if any, of the candidates— $A$ ,  $B$ ,  $C$ , or  $D$ —is not eliminated by the *necessary* condition test as a necessary condition for target feature  $G$ ?
- Which, if any, of the candidates— $A$ ,  $B$ ,  $C$ , or  $D$ —is not eliminated by *either* test?

EXAMPLE:

Case 1:	$A$	$B$	$\sim C$	$D$	$\sim G$
Case 2:	$\sim A$	$B$	$C$	$D$	$G$
Case 3:	$A$	$\sim B$	$C$	$D$	$G$

- Only  $C$  passes the SCT.
- Only  $C$  and  $D$  pass the NCT.
- Only  $C$  passes both tests.

1.	Case 1:	$A$	$B$	$C$	$D$	$G$
	Case 2:	$\sim A$	$B$	$\sim C$	$D$	$\sim G$
	Case 3:	$A$	$\sim B$	$C$	$\sim D$	$G$
2.	Case 1:	$A$	$B$	$C$	$D$	$G$
	Case 2:	$\sim A$	$B$	$C$	$D$	$G$
	Case 3:	$A$	$\sim B$	$C$	$D$	$G$
3.	Case 1:	$A$	$B$	$C$	$D$	$\sim G$
	Case 2:	$\sim A$	$B$	$C$	$D$	$G$
	Case 3:	$A$	$\sim B$	$C$	$\sim D$	$G$
4.	Case 1:	$A$	$B$	$C$	$D$	$G$
	Case 2:	$\sim A$	$\sim B$	$C$	$D$	$G$
	Case 3:	$A$	$B$	$\sim C$	$\sim D$	$\sim G$
5.	Case 1:	$A$	$B$	$C$	$D$	$\sim G$
	Case 2:	$\sim A$	$B$	$C$	$\sim D$	$\sim G$
	Case 3:	$A$	$\sim B$	$\sim C$	$D$	$G$
6.	Case 1:	$A$	$B$	$\sim C$	$D$	$G$
	Case 2:	$\sim A$	$\sim B$	$C$	$D$	$\sim G$
	Case 3:	$A$	$\sim B$	$C$	$\sim D$	$\sim G$
7.	Case 1:	$A$	$B$	$\sim C$	$D$	$\sim G$
	Case 2:	$\sim A$	$B$	$\sim C$	$D$	$\sim G$
	Case 3:	$A$	$B$	$\sim C$	$\sim D$	$\sim G$
8.	Case 1:	$A$	$B$	$C$	$D$	$G$
	Case 2:	$\sim A$	$\sim B$	$C$	$D$	$G$
	Case 3:	$A$	$\sim B$	$\sim C$	$D$	$\sim G$

**EXERCISE IV**

Imagine that your desktop computer system won't work, and you want to find out why. After checking to make sure that it is plugged in, you experiment with a new central processing unit (CPU), a new monitor (MO), and new system software (SW) in the combinations on the table below. The candidates for necessary conditions and sufficient conditions of failure are the plug position (in or out), the CPU (old or new), the monitor (old or new), and the software (old or new). For each candidate, say (1) which cases, if any, eliminate it as a sufficient condition of your computer's failure and (2) which cases, if any, eliminate it as a necessary condition of your computer's failure. Which candidates, if any, are not eliminated as a sufficient condition of failure? As a necessary condition of failure? Does it follow that these candidates are necessary conditions or sufficient conditions of failure? Why or why not?

(continued)

	<b>Plug</b>	<b>CPU</b>	<b>Monitor</b>	<b>Software</b>	<b>Result</b>
Case 1	In	Old CPU	Old MO	Old SW	Works
Case 2	In	Old CPU	Old MO	New SW	Works
Case 3	In	Old CPU	New MO	Old SW	Fails
Case 4	In	Old CPU	New MO	New SW	Works
Case 5	In	Old CPU	Old MO	Old SW	Works
Case 6	In	Old CPU	Old MO	New SW	Works
Case 7	In	Old CPU	New MO	Old SW	Fails
Case 8	In	Old CPU	New MO	New SW	Works
Case 9	In	New CPU	Old MO	Old SW	Fails
Case 10	In	New CPU	Old MO	New SW	Works
Case 11	In	New CPU	New MO	Old SW	Fails
Case 12	In	New CPU	New MO	New SW	Works

**EXERCISE V**

After a banquet, several diners get sick and die. You suspect that something they ate or drank caused their deaths. The following table records their meals and fates. The target feature is death. The candidates for necessary conditions and sufficient conditions of death are the soup, entrée, wine, and dessert. For each candidate, say (1) which cases, if any, eliminate it as a sufficient condition of death and (2) which cases, if any, eliminate it as a necessary condition of death. Which candidates, if any, are not eliminated as a sufficient condition of death? Which candidates, if any, are not eliminated as a necessary condition of death? Does it follow that these candidates are necessary conditions or sufficient conditions of death? Why or why not?

<b>Diners</b>	<b>Soup</b>	<b>Entrée</b>	<b>Wine</b>	<b>Dessert</b>	<b>Result</b>
Ann	Tomato	Chicken	White	Pie	Alive
Barney	Tomato	Fish	Red	Cake	Alive
Cathy	Tomato	Beef	Red	Ice Cream	Dead
Doug	Tomato	Beef	Red	Cake	Alive
Emily	Tomato	Fish	Red	Pie	Alive
Fred	Tomato	Fish	Red	Cake	Alive
Gertrude	Leek	Fish	White	Pie	Alive
Harold	Tomato	Beef	White	Cake	Alive
Irma	Leek	Fish	Red	Pie	Alive
Jack	Leek	Beef	Red	Ice Cream	Dead
Ken	Leek	Chicken	Red	Ice Cream	Alive
Leslie	Tomato	Chicken	White	Cake	Alive

## RIGOROUS TESTING

Going back to Table 1, it is easy to see that candidates  $A$ ,  $B$ ,  $C$ , and  $D$  are not eliminated by the NCT as necessary conditions of target  $G$ , as  $G$  is present in only one case (Case 1) and  $A$ ,  $B$ ,  $C$ , and  $D$  are present there as well. So far, so good. But if we wanted to test these features more rigorously, it would be important to find more cases in which target  $G$  was present and see whether these candidates are also present and thus continue to survive the NCT.

The following table gives a more extreme example of nonrigorous testing:

TABLE 3

Case 1:	$A$	$\sim B$	$C$	$D$	$G$
Case 2:	$A$	$\sim B$	$\sim C$	$\sim D$	$\sim G$
Case 3:	$A$	$\sim B$	$C$	$\sim D$	$\sim G$
Case 4:	$A$	$\sim B$	$\sim C$	$D$	$G$

Here candidate feature  $A$  is eliminated by SCT (in Cases 2 and 3) but is not eliminated by NCT, so it is a possible necessary condition but not a possible sufficient condition for target feature  $G$ .  $B$  is not eliminated by SCT but is eliminated by NCT (in Cases 1 and 4), so it is a possible sufficient condition but not a possible necessary condition for target feature  $G$ .  $C$  is eliminated by both rules (in Cases 3 and 4). Only  $D$  is not eliminated by either test, so it is the only candidate for being both a necessary and a sufficient condition for  $G$ .

The peculiarity of this example is that candidate  $A$  is always present whether target  $G$  is present or not, and candidate  $B$  is always absent whether target  $G$  is absent or not. Now if something is always present, as  $A$  is, then it cannot possibly fail the NCT; for there cannot be a case where the target is present and the candidate is absent if the candidate is *always* present. If we want to test candidate  $A$  rigorously under the NCT, then we should try to find cases in which  $A$  is absent and then check to see whether  $G$  is absent as well.

In reverse fashion, but for similar reasons, if we want to test candidate  $B$  rigorously under the SCT, then we should try to find cases in which  $B$  is present and then check to see if  $G$  is present as well. If we restrict our attention to cases where  $B$  is always absent, as in Table 3, then  $B$  cannot possibly fail the SCT, but passing that test will be trivial for  $B$  and so will not even begin to show that  $B$  is a sufficient condition for  $G$ .

Now consider two more sets of data just like Table 2, except with regard to the target feature,  $G$ :

TABLE 4

Case 1:	$A$	$B$	$C$	$D$	$G$
Case 2:	$\sim A$	$B$	$C$	$D$	$G$
Case 3:	$A$	$\sim B$	$C$	$\sim D$	$G$

TABLE 5

Case 1:	$A$	$B$	$C$	$D$	$\sim G$
Case 2:	$\sim A$	$B$	$C$	$D$	$\sim G$
Case 3:	$A$	$\sim B$	$C$	$\sim D$	$\sim G$

Because  $G$  is present in all of the cases in Table 4, no candidate can be eliminated by the SCT as a sufficient condition for target feature  $G$ . This result is trivial, however. Table 4 does not provide rigorous testing for a sufficient condition of  $G$ , because our attention is restricted to a range of cases that is too narrow. Nothing could possibly be eliminated as a sufficient condition of  $G$  as long as  $G$  is always present.

Similarly,  $G$  is absent in all of the cases in Table 5, so no candidate can be eliminated by the NCT as a necessary condition of target feature  $G$ . Still, because this data is so limited, its failure to eliminate candidates does not even begin to show that anything is a necessary condition of  $G$ .

For both rules, then, rigorous testing involves seeking out cases in which failing the test is a live possibility. For the SCT, this requires looking both at cases in which the candidates are present and also at cases in which the target is absent. For the NCT, rigorous testing requires looking both at cases in which the candidates are absent and also at cases in which the target is present. Without cases like these, passing the tests is rather like a person bragging that he has never struck out when, in fact, he has never come up to bat.

## REACHING POSITIVE CONCLUSIONS

Suppose that we performed rigorous testing on candidate  $C$ , and it passed the SCT with flying colors. Can we now draw the positive conclusion that  $C$  is a sufficient condition for the target  $G$ ? That depends on which kinds of candidates and cases have been considered. Since rigorous testing was passed, these three conditions are met:

1. We have tested some cases in which the candidate,  $C$ , is present.
2. We have tested some cases in which the target,  $G$ , is absent.
3. We have not found any case in which the candidate,  $C$ , is present and the target,  $G$ , is absent.

In cases that meet these three conditions, we sometimes face a dilemma. More than one candidate might pass this rigorous testing. It is possible that both of these candidates is sufficient for the target feature, but there is often some reason to worry that only one of them is really causing the effect. In order to test this hypothesis, we can add another restriction:

4. If there is any other candidate,  $D$ , that is never present where the target,  $G$ , is absent, then we have tested cases where  $C$  is present and  $D$  is absent.

The target,  $G$ , must also be present in these cases, since  $C$  is present and Condition 3 has already been met. Testing this group of cases can reassure us that it is not only candidate  $D$  that is sufficient for the target,  $G$ .

Finally, for it to be reasonable to reach a positive conclusion that  $C$  is sufficient for  $G$ , this further condition must also be met:

5. We have tested enough cases of the various kinds that are likely to include a case in which  $C$  is present and  $G$  is absent if there is any such case.

This new condition cannot be applied in the mechanical way that conditions 1–4 could be applied. To determine whether condition 5 is met, we need to rely on *background information* about how many cases are “enough” and about which kinds of cases “are likely to include a case in which  $C$  is present and  $G$  is absent, if there is any such case.” For example, if we are trying to figure out whether our new software is causing our computer to crash, we do not need to try the same kind of computer in different colors. What we need to try are different kinds of CPUs, monitors, software, and so on, because we know that these are the kinds of factors that can affect performance. Background information like this is what tells us when we have tested enough cases of the right kinds.

Of course, our background assumptions might turn out to be wrong. Even if we have tested many variations of every feature that we think might be relevant, we still might be surprised and find a further case in which  $C$  and  $\sim G$  are present. All that shows, however, is that our inference is defeasible, like all inductive arguments. Despite the possibility that future discoveries might undermine it, our inductive inference can still be strong if our background beliefs are justified and if we have looked long and hard without finding any case in which  $C$  is present and  $G$  is absent.

Similar rules apply in reverse to positive conclusions about necessary conditions. We have good reason to suppose that candidate  $C$  is a necessary condition for target  $G$ , if the following conditions are met:

1. We have tested some cases in which the candidate,  $C$ , is absent.
2. We have tested some cases in which the target,  $G$ , is present.
3. We have not found any case in which the candidate,  $C$ , is absent and the target,  $G$ , is present.
4. If there is any other candidate,  $D$ , that is never absent where the target,  $G$ , is present, then we have tested cases where  $C$  is absent and  $D$  is present.
5. We have tested enough cases of the various kinds that are likely to include a case in which  $C$  is absent and  $G$  is present, if there is any such case.

This argument again depends on background assumptions in determining whether condition 5 is met. This argument is also defeasible, as before.

Nonetheless, if our background assumptions are justified, the fact that conditions 1–5 are all met can still provide a strong reason for the positive conclusion that candidate C is a necessary condition for target G.

The SCT and NCT themselves are still negative and deductive; but that does not make them better than the positive tests encapsulated in conditions 1–5. The negative SCT and NCT are of no use when we need to argue that some condition *is* sufficient or *is* necessary. Such positive conclusions can be reached only by applying something like condition 5, which will require background information. These inductive arguments might not be as clear-cut or secure as the negative ones, but they can still be inductively strong under the right circumstances. That is all they claim to be.

## APPLYING THESE METHODS TO FIND CAUSES

In stating the SCT and NCT and applying these tests to abstract patterns of conditions to eliminate candidates, our procedure was fairly mechanical. We cannot be so mechanical when we try to reach positive conclusions that certain conditions are necessary, sufficient, or both. Applying these rules to actual concrete situations introduces a number of further complications, especially when using our tests to determine causes.

### NORMALITY

First, it is important to keep in mind that, in our ordinary understanding of causal conditions, we usually take it for granted that the setting is normal. It is part of common knowledge that if you strike a match, then it will light. Thus, we consider striking a match sufficient to make it light. But if someone has filled the room with carbon dioxide, then the match will not light, no matter how it is struck. Here one may be inclined to say that, after all, striking a match is not sufficient to light it. We might try to be more careful and say that if a match is struck *and* the room is not filled with carbon dioxide, then it will light. But this new conditional overlooks other possibilities—for example, that the room has been filled with nitrogen, that the match has been fireproofed, that the wrong end of the match was struck, that the match has already been lit, and so forth. It now seems that the antecedent of our conditional will have to be endlessly long in order to specify a true or genuine sufficient condition. In fact, however, we usually feel quite happy with saying that if you strike a match, then it will light. We simply do not worry about the possibility that the room has been filled with carbon dioxide, the match has been fireproofed, and so on. Normally we think that things are normal, and give up this assumption only when some good reason appears for doing so.

These reflections suggest the following *contextualized* restatement of our original definitions of sufficient conditions and necessary conditions:

*F* is a sufficient condition for *G* if and only if, whenever *F* is present in a normal context, *G* is present there as well.

*F* is a necessary condition for *G* if and only if, whenever *F* is absent from a normal context, *G* is absent from it as well.

What will count as a normal context will vary with the type and the aim of an investigation, but all investigations into causally sufficient conditions and causally necessary conditions take place against the background of many factors that are taken as fixed.

### BACKGROUND ASSUMPTIONS

If we are going to subject a causal hypothesis to rigorous testing with the SCT and the NCT, we have to seek out a wide range of cases that might refute that hypothesis. In general, the wider the range of possible refuters the better. Still, some limit must be put on this activity or else testing will get hopelessly bogged down. If we are testing a drug to see whether it will cure a disease, we should try it on a variety of people of various ages, medical histories, body types, and so on, but we will not check to see whether it works on people named Edmund or check to see whether it works on people who drive Volvos. Such factors, we want to say, are plainly irrelevant. But what makes them irrelevant? How do we distinguish relevant from irrelevant considerations?

The answer to this question is that our reasoning about causes occurs within a framework of beliefs that we take to be established as true. This framework contains a great deal of what is called *common knowledge*—knowledge we expect almost every sane adult to possess. We all know, for example, that human beings cannot breathe underwater, cannot walk through walls, cannot be in two places at once, and so on. The stock of these commonplace beliefs is almost endless. Because they are commonplace beliefs, they tend not to be mentioned; yet, they play an important role in distinguishing relevant factors from irrelevant ones.

*Specialized* knowledge also contains its own principles that are largely taken for granted by experts. Doctors, for example, know a great deal about the detailed structure of the human body, and this background knowledge constantly guides their thought in dealing with specific illnesses. Even if someone claimed to discover that blood does not circulate, no doctor would take the time to refute that claim.

It might seem close-minded to refuse to consider a possibility that someone else suggests. However, giving up our basic beliefs can be very costly. A doctor who took seriously the suggestion that blood does not circulate, for example, would have to abandon our whole way of viewing humans and

other animals, along with the rest of biology and science. It is not clear how this doctor could go on practicing medicine. Moreover, there is usually no practical alternative in real life. When faced with time pressure and limited information, we have no way to judge new ideas without taking some background assumptions for granted.

### A DETAILED EXAMPLE

To get a clearer idea of the complex interplay between our tests and the reliance on background information, it will be helpful to look in some detail at actual applications of these tests. For this purpose, we will examine an attempt to find the cause of a particular phenomenon, an outbreak of what came to be known as Legionnaires' disease. The example not only shows how causal reasoning relies on background assumptions, it has another interesting feature as well: In the process of discovering the cause of Legionnaires' disease, the investigators were forced to *abandon* what was previously taken to be a well-established causal generalization. In fact, until it was discarded, this false background principle gave them no end of trouble.

The story began at an otherwise boring convention:

The 58th convention of the American Legion's Pennsylvania Department was held at the Bellevue-Stratford Hotel in Philadelphia from July 21 through 24, 1976. . . . Between July 22 and August 3, 149 of the conventioners developed what appeared to be the same puzzling illness, characterized by fever, coughing and pneumonia. This, however, was an unusual, explosive outbreak of pneumonia with no apparent cause. . . . Legionnaires' disease, as the illness was quickly named by the press, was to prove a formidable challenge to epidemiologists and laboratory investigators alike.<sup>2</sup>

Notice that at this stage the researchers begin with the assumption that they are dealing with a single illness and not a collection of similar but different illnesses. That assumption could turn out to be wrong; but, if the symptoms of the various patients are sufficiently similar, this is a natural starting assumption. Another reasonable starting assumption is that this illness had a single causative agent. This assumption, too, could turn out to be false, though it did not. The assumption that they were dealing with a single disease with a single cause was at least a good simplifying assumption, one to be held onto until there was good reason to give it up. In any case, we now have a clear specification of our target feature, G: the occurrence of a carefully described illness that came to be known as Legionnaires' disease. The situation concerning it was puzzling because people had contracted a disease with symptoms much like those of pneumonia, yet they had not tested positive for any of the known agents that cause such diseases.

The narrative continues as follows:

The initial step in the investigation of any epidemic is to determine the character of the illness, who has become ill and just where and when. The next step is to

find out what was unique about the people who became ill: where they were and what they did that was different from other people who stayed well. Knowing such things may indicate how the disease agent was spread and thereby suggest the identity of the agent and where it came from.<sup>3</sup>

Part of this procedure involves a straightforward application of the NCT: Was there any interesting feature that was always present in the history of people who came down with the illness? Progress was made almost at once on this front:

We quickly learned that the illness was not confined to Legionnaires. An additional 72 cases were discovered among people who had not been directly associated with the convention. They had one thing in common with the sick conventioneers: for one reason or another they had been in or near the Bellevue-Stratford Hotel.<sup>4</sup>

Strictly speaking, of course, all these people who had contracted the disease had more than one thing in common. They were, for example, all alive at the time they were in Philadelphia, and being alive is, in fact, a necessary condition for getting Legionnaires' disease. But the researchers were not interested in this necessary condition because it is a normal background condition for the contraction of any disease. Furthermore, it did not provide a condition that distinguished those who contracted the disease from those who did not. The overwhelming majority of people who were alive at the time did not contract Legionnaires' disease. Thus, the researchers were not interested in this necessary condition because it would fail so badly when tested by the SCT as a sufficient condition. On the basis of common knowledge and specialized medical knowledge, a great many other conditions were also kept off the candidate list.

One prime candidate on the list was presence at the Bellevue-Stratford Hotel. The application of the NCT to this candidate was straightforward. Everyone who had contracted the disease had spent time in or near that hotel. Thus, presence at the Bellevue-Stratford could not be eliminated as a necessary condition of Legionnaires' disease.

The application of the SCT was more complicated, because not everyone who stayed at the Bellevue-Stratford contracted the disease. Other factors made a difference: "Older conventioneers had been affected at a higher rate than younger ones, men at three times the rate for women." Since some young women (among others) who were present at the Bellevue-Stratford did not get Legionnaires' disease, presence at that hotel could be eliminated as a sufficient condition of Legionnaires' disease. Nonetheless, it is part of medical background knowledge that susceptibility to disease often varies with age and gender. Given these differences, some people who spent time at the Bellevue-Stratford were at higher risk for contracting the disease than others. The investigation so far suggested that, for some people, being at the Bellevue-Stratford was connected with a sufficient condition for contracting Legionnaires' disease. Indeed, the conjunction of spending time at the

Bellevue-Stratford and being susceptible to the disease could not be ruled out by the SCT as a sufficient condition of getting the disease.

As soon as spending time at the Bellevue-Stratford became the focus of attention, other hypotheses naturally suggested themselves. Food poisoning was a reasonable suggestion, since it is part of medical knowledge that diseases are sometimes spread by food. It was put on the list of possible candidates, but failed. Investigators checked each local restaurant and each function where food and drink were served. Some of the people who ate in each place did not get Legionnaires' disease, so the food at these locations was eliminated by the SCT as a sufficient condition of Legionnaires' disease. These candidates were also eliminated by the NCT as necessary conditions because some people who did get Legionnaires' disease did not eat at each of these restaurants and functions. Thus, the food and drink could not be the cause.

Further investigation turned up another important clue to the cause of the illness.

Certain observations suggested that the disease might have been spread through the air. Legionnaires who became ill had spent on the average about 60 percent more time in the lobby of the Bellevue-Stratford than those who remained well; the sick Legionnaires had also spent more time on the sidewalk in front of the hotel than their unaffected fellow conventioners. . . . It appeared, therefore, that the most likely mode of transmission was airborne.<sup>5</sup>

Merely breathing air in the lobby of the Bellevue-Stratford Hotel still could not be a necessary or sufficient condition, but the investigators reasoned that something in the lobby air probably caused Legionnaires' disease, since the rate of the disease varied up or down in proportion to the time spent in the lobby (or near it on the sidewalk in front). This is an application of the method of concomitant variation, which will be discussed soon.

Now that the focus was on the lobby air, the next step was to pinpoint a specific cause in that air. Again appealing to background medical knowledge, there seemed to be three main candidates for the airborne agents that could have caused the illness: "heavy metals, toxic organic substances, and infectious organisms." Examination of tissues taken from patients who had died from the disease revealed "no unusual levels of metallic or toxic organic substances that might be related to the epidemic," so this left an infectious organism as the remaining candidate. Once more we have an application of NCT. If the disease had been caused by heavy metals or toxic organic substances, then there would have been unusually high levels of these substances in the tissues of those who had contracted the disease. Because this was not always so, these candidates were eliminated as necessary conditions of the disease.

Appealing to background knowledge once more, it seemed that a bacterium would be the most likely source of an airborne disease with the symptoms of Legionnaires' disease. But researchers had already made a routine

check for bacteria that cause pneumonia-like diseases, and they had found none. For this reason, attention was directed to the possibility that some previously unknown organism had been responsible but had somehow escaped detection.

It turned out that an undetected and previously unknown bacterium *had* caused the illness, but it took more than four months to find this out. The difficulties encountered in this effort show another important fact about the reliance on a background assumption: Sometimes it turns out to be false. To simplify, the standard way to test for the presence of bacteria is to try to grow them in culture dishes—flat dishes containing nutrients that bacteria can live on. If, after a reasonable number of tries, a colony of a particular kind of bacterium does not appear, then it is concluded that the bacterium is not present. As it turned out, the bacterium that caused Legionnaires' disease would not grow in the cultures commonly used to detect the presence of bacteria. Thus, an important background assumption turned out to be false.

After a great deal of work, a suspicious bacterium was detected using a live-tissue culture rather than the standard synthetic culture. The task, then, was to show that this particular bacterium in fact caused the disease. Again to simplify, when people are infected by a particular organism, they often develop antibodies that are specifically aimed at this organism. In the case of Legionnaires' disease, these antibodies were easier to detect than the bacterium itself. They also remained in the patients' bodies after the infection had run its course. We thus have another chance to apply the NCT: If Legionnaires' disease was caused by this particular bacterium, then whenever the disease was present, this antibody should be present as well. The suspicious bacterium passed this test with flying colors and was named, appropriately enough, *Legionella pneumophila*. Because the investigators had worked so hard to test such a wide variety of candidates, they assumed that the disease must have some cause among the candidates that they checked. So, since only one candidate remained, they felt justified in reaching a positive conclusion that the bacterium was a necessary condition of Legionnaires' disease.

The story of the search for the cause of Legionnaires' disease brings out two important features of the use of inductive methods in the sciences. First, it involves a complicated interplay between what is already established and what is being tested. Confronted with a new problem, established principles can be used to suggest theoretically significant hypotheses to be tested. The tests then eliminate some hypotheses and leave others. If, at the end of the investigation, a survivor remains that fits in well with our previously established principles, then the stock of established principles is increased. The second thing that this example shows is that the inductive method is fallible. Without the background of established principles, the application of inductive principles like the NCT and the SCT would be undirected; yet sometimes these established principles let us down, for they can turn out to

be false. The discovery of the false background principle that hindered the search for the cause of Legionnaires' disease led to important revisions in laboratory techniques. The discovery that certain fundamental background principles are false can lead to revolutionary changes in science.

### DISCUSSION QUESTIONS

1. Sometimes we describe necessary conditions as causes, but at other times we describe sufficient conditions as causes. Why? Be sure to give at least two different examples of each pattern.
2. *Legionella pneumophila* is necessary for Legionnaire's disease, but so are being alive, having blood, and so on. Why do we think that Legionnaire's disease is caused by *Legionella pneumophila* instead of being caused by being alive, having blood, and other necessary conditions?

## CONCOMITANT VARIATION

The use of the sufficient condition test and the necessary condition test depends on certain features of the world being sometimes present and sometimes absent. Some features of the world, however, are always present to some degree. Because they are always present, the NCT will never eliminate them as possible necessary conditions of any event, and the SCT will never eliminate anything as a sufficient condition for them. Yet, the *extent or degree* to which a feature exists in the world is often a significant phenomenon that demands causal explanation.

An example should make this clear. In recent decades, a controversy has raged over the impact of acid rain on the environment of the northeastern United States and Canada. Part of the controversy involves the proper interpretation of the data that have been collected. The controversy has arisen for the following reason: The atmosphere always contains a certain amount of acid, much of it from natural sources. It is also known that an excess of acid in the environment can have severe effects on both plants and animals. Lakes are particularly vulnerable to the effects of acid rain. Finally, it is also acknowledged that industries, mostly in the Midwest, discharge large quantities of sulfur dioxide ( $\text{SO}_2$ ) into the air, and this increases the acidity of water in the atmosphere. The question—and here the controversy begins—is whether the contribution of acid from these industries is the cause of the environmental damage downwind of them.

How can we settle such a dispute? The two rules we have introduced provide no immediate help, for, as we have seen, they provide a rigorous test of a causal hypothesis only when we can find contrasting cases with the presence or the absence of a given feature. The NCT provides a rigorous

test for a necessary condition only if we can find cases in which the feature does not occur and then check to make sure that the target feature does not occur either. The SCT provides a rigorous test for a sufficient condition only when we can find cases in which the target phenomenon is absent and then check whether the candidate sufficient condition is absent as well. In this case, however, neither check applies, for there is always a certain amount of acid in the atmosphere, so it is not possible to check what happens when atmospheric acid is completely absent. Similarly, environmental damage, which is the target phenomenon under investigation, is so widespread in our modern industrial society that it is also hard to find a case in which it is completely absent.

So, if there is always acid in the atmosphere, and environmental damage always exists at least to some extent, how can we determine whether the SO<sub>2</sub> released into the atmosphere is *significantly* responsible for the environmental damage in the affected areas? Here we use what John Stuart Mill called the *Method of Concomitant Variation*. We ask whether the amount of environmental damage varies directly in proportion to the amount of SO<sub>2</sub> released into the environment. If environmental damage increases with the amount of SO<sub>2</sub> released into the environment and drops when the amount of SO<sub>2</sub> is lowered, this means that the level of SO<sub>2</sub> in the atmosphere is *positively correlated* with environmental damage. We would then have good reason to believe that lowering SO<sub>2</sub> emissions would lower the level of environmental damage, at least to some extent.

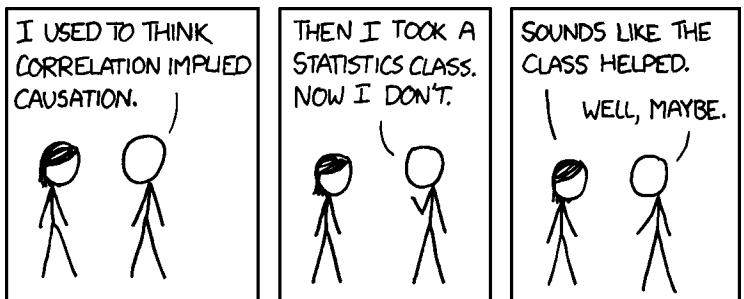
Arguments relying on the method of concomitant variation are difficult to evaluate, especially when there is no generally accepted background theory that makes sense of the concomitant variation. Some such variations are well understood. For example, most people know that the faster you drive, the more gasoline you consume. (Gasoline consumption varies *directly* with speed.) Why? There is a good theory here: It takes more energy to drive at a high speed than at a low speed, and this energy is derived from the gasoline consumed in the car's engine. Other correlations are less well understood. Reportedly, there seems to be a correlation between how much a woman smokes during pregnancy and how happy her children are when they reach age thirty. The correlation here is not nearly as good as the correlation between gasoline consumption and speed, for many people are very happy at age thirty even though their mothers smoked a lot during pregnancy, and many others are very unhappy at age thirty even though their mothers never smoked. Furthermore, no generally accepted background theory has been found to explain the correlation that does exist.

This reference to background theory is important, because two sets of phenomena can be correlated to a very high degree, even with no direct causal relationship between them. A favorite example that appears in many statistics texts is the discovered positive correlation in boys between foot size and quality of handwriting. It is hard to imagine a causal relation holding in either direction. Having big feet should not make you write better

and, just as obviously, writing well should not give you big feet. The correct explanation is that both foot size and handwriting ability are positively correlated with age. Here, a noncausal correlation between two phenomena (foot size and handwriting ability) is explained by a third common correlation (maturation) that *is* causal.

At times, it is possible to get causal correlations *backward*. For example, a few years ago, sports statisticians discovered a negative correlation between forward passes thrown and winning in football. That is, the more forward passes a team threw, the less chance it had of winning. This suggested that passing is not a good strategy, since the more you do it, the more likely you are to lose. Closer examination showed, however, that the causal relationship, in fact, went in the other direction. Toward the end of a game, losing teams tend to throw a great many passes in an effort to catch up. In other words, teams throw a lot of passes because they are losing, rather than the other way around.

Finally, some correlations seem inexplicable. For example, a strong positive correlation reportedly holds between the birth rate in Holland and the number of storks nesting in chimneys. There is, of course, a background theory that would explain this—storks bring babies—but that theory is not favored by modern science. For the lack of any better background theory, the phenomenon just seems weird.



Courtesy of Randall Munroe

So, given a strong correlation between phenomena of types *A* and *B*, four possibilities exist:

1. *A* is the cause of *B*.
2. *B* is the cause of *A*.
3. Some third thing is the cause of both.
4. The correlation is simply accidental.

Before we accept any one of these possibilities, we must have good reasons for preferring it over the other three.

One way to produce such a reason is to manipulate *A* or *B*. If we vary factor *A* up and down, but *B* does not change at all, this finding provides some reason against possibility 1, since *B* would normally change along with *A* if *A* did cause *B*. Similarly, if we manipulate *B* up and down, but *A* does not vary at all, this result provides some reason against alternative 2 and for

the hypothesis that *B* does not cause *A*. Together these manipulations can reduce the live options to items 3 and 4.

Many scientific experiments work this way. When scientists first discovered the correlation between smoking and lung cancer, some cigarette manufacturers responded that lung cancer might cause the desire to smoke or there might be a third cause of both smoking and lung cancer that explains the correlation. Possibly, it was suggested, smoking relieves discomfort due to early lung cancer or due to a third factor that itself causes lung cancer. To test these hypotheses, scientists manipulated the amount of smoking by lab animals. When all other factors were held as constant as possible, but smoking was increased, lung cancer increased; and when smoking went down, lung cancer went down. These results would not have occurred if some third factor had caused both smoking and lung cancer but remained stable as smoking was manipulated. The findings would also have been different if incipient lung cancer caused smoking, but had remained constant as scientists manipulated smoking levels. Such experiments can, thus, help us rule out at least some of the options 1–4.

Direct manipulation like this is not always possible or ethically permissible. The data would probably be more reliable if the test subjects were human beings rather than lab animals, but that is not an ethical option. Perhaps more complicated statistical methods could produce more reliable results, but they often require large amounts and special kinds of data. Such data is, unfortunately, often unavailable.

#### EXERCISE VI

In each of the following examples a strong correlation, either negative or positive, holds between two sets of phenomena, *A* and *B*. Try to decide whether *A* is the cause of *B*, *B* is the cause of *A*, both are caused by some third factor, *C*, or the correlation is simply accidental. Explain your choice.

1. For a particular United States president, there is a negative correlation between the number of hairs on his head (*A*) and the population of China (*B*).
2. My son's height (*A*) increases along with the height of the tree outside my front door (*B*).
3. It has been claimed that there is a strong positive correlation between those students who take sex education courses (*A*) and those who contract venereal disease (*B*).
4. At one time there was a strong negative correlation between the number of mules in a state (*A*) and the salaries paid to professors at the state university (*B*). In other words, the more mules, the lower professional salaries.<sup>6</sup>
5. There is a high positive correlation between the number of fire engines in a particular borough in New York City (*A*) and the number of fires that occur there (*B*).<sup>7</sup>

6. "Washington (UPI)—Rural Americans with locked doors, watchdogs or guns may face as much risk of burglary as neighbors who leave doors unlocked, a federally financed study says. The study, financed in part by a three-year \$170,000 grant from the Law Enforcement Assistance Administration, was based on a survey of nearly 900 families in rural Ohio. Sixty percent of the rural residents surveyed regularly locked doors [A], but were burglarized more often than residents who left doors unlocked [B]."<sup>8</sup>
7. The speed of a car (A) is exactly the same as the speed of its shadow (B).
8. The length of a runner's ring finger minus the length of the runner's index finger (A) is correlated with the runner's speed in the one-hundred-yard dash (B).

### DISCUSSION QUESTIONS

1. After it became beyond doubt that smoking is dangerous to people's health, a new debate arose concerning the possible health hazards of secondhand smoke on nonsmokers. Collect statements pro and con on this issue and evaluate the strength of the inductive arguments on each side.
2. The high positive correlation between CO<sub>2</sub> concentrations in the atmosphere and the Earth's mean surface temperatures is often cited as evidence that increases in atmospheric CO<sub>2</sub> cause global warming. This argument is illustrated by the famous "hockey stick" diagram in Al Gore's *An Inconvenient Truth*. Is this argument persuasive? How could skeptics about global warming respond?

### NOTES

<sup>1</sup> Mill's "methods of experimental inquiry" are found in book 3, chap. 8 of his *A System of Logic* (London: John W. Parker, 1843). Mill's method of difference, method of agreement, and joint method parallel our SCT, NCT, and Joint Test, respectively. Our simplification of Mill's methods derives from Brian Skyrms, *Choice and Chance*, 3rd ed. (Belmont, CA: Wadsworth, 1986), chap. 4.

<sup>2</sup> These excerpts are drawn from David W. Fraser and Joseph E. McDade, "Legionellosis," *Scientific American*, October 1979, 82–99.

<sup>3</sup> Ibid.

<sup>4</sup> Ibid.

<sup>5</sup> Ibid.

<sup>6</sup> From Gregory A. Kimble, *How to Use (and Misuse) Statistics* (Englewood Cliffs, NJ: Prentice-Hall, 1978), 182.

<sup>7</sup> Ibid.

<sup>8</sup> "Locked Doors No Bar to Crime, Study Says," *Santa Barbara [California] Newspress*, Wednesday, February 16, 1977. This title suggests that locking your doors will not increase safety. Is that a reasonable lesson to draw from this study?

## CHANCES

The kinds of arguments discussed in the preceding three chapters are inductive, so they need not meet the deductive standard of validity. They are, instead, intended to meet the inductive standard of strength. Whereas deductive validity hinges on what is possible, inductive strength hinges on what is probable. Roughly, an argument is inductively strong to the extent that its premises make its conclusion more likely or probable. Hence, just as we can get a better theoretical understanding of deductive validity by studying formal logic, as we did in Chapters 6 and 7, so we can get a better theoretical understanding of inductive strength by studying probability, as we will do in this chapter. To complete our survey of inductive arguments, this chapter offers an elementary discussion of probability. It begins by illustrating several common mistakes about probability. To help avoid these fallacies, we need to approach probability more carefully, so formal laws of probability are presented along with Bayes's theorem.

### SOME FALLACIES OF PROBABILITY

Probability is pervasive. We all assume or make probability judgments throughout our lives. We do so whenever we form a belief about which we are not certain, as in all of the kinds of inductive arguments studied in Chapters 8–10. Such arguments do not pretend to reach their conclusions with certainty, even if their premises are true. They merely try to show that a conclusion is likely or probable. Judgments about probability are, thus, assumed in assessing such arguments and beliefs. Probability also plays a crucial role in our most important decisions. Mistakes about probability can then lead to disasters. Doctors lose patients' lives, stockbrokers lose clients' money, and coaches lose games because they overestimate or underestimate probabilities. Such mistakes are common and fall into several regular patterns. It is useful to understand these fallacies, so that we can learn to avoid them.

#### THE GAMBLER'S FALLACY

Casinos thrive partly because so many gamblers misunderstand probability. One mistake is so common that it has been dubbed *the gambler's fallacy*. When people have a run of bad luck, they often increase their bets because they

assume that they are due for a run of good luck to even things out. Gambling systems are sometimes based on this fallacious idea. People keep track of the numbers that come up on a roulette wheel, trying to discover a number that has not come up for a long time. They then pile their money on that number on the assumption that it is due. They usually end up losing a bundle.

These gamblers seem to assume, "In the long run, things will even out (or average out)." Interpreted one way, this amounts to what mathematicians call *the law of large numbers*, and it is perfectly correct. When flipping a coin, we expect it to come up heads half the time, so it should come up heads five times in ten flips. If we actually check this out, however, we discover that the number of times it comes up heads in ten flips varies significantly from this predicted value—sometimes coming up heads more than five times, sometimes coming up fewer. What the law of large numbers tells us is that the actual percentage of heads will tend to come closer to the theoretically predicted percentage of heads the more trials we make. If you flipped a coin a million times, it would be very surprising if the percentage of heads were more than 1 percent away from the predicted 50 percent.



This law of large numbers is often misunderstood in a way that leads to the gambler's fallacy. Some people assume that each possible outcome will occur the average number of times in each series of trials. To see that this is a fallacy,

we can go back to flipping coins. Toss a coin until it comes up heads three times in a row. (This will take less time than you might imagine.) What is the probability that it will come up heads a fourth time? Put crudely, some people think that the probability of it coming up heads again must be very small, because it is unlikely that a fair coin will come up heads four times in a row, so a tails is needed to even things out. That is wrong. The chances of getting heads on any given toss are the same, regardless of what happened on previous tosses. Previous results cannot affect the probabilities on this new toss.

## HEURISTICS

In daily life, we often have to make decisions quickly without full information. To deal with this overload of decisions, we commonly employ what cognitive psychologists call *heuristics*. Technically, a heuristic is a general strategy for solving a problem or coming to a decision. For example, a good heuristic for solving geometry problems is to start with the conclusion you are trying to reach and then work backward.

Recent research in cognitive psychology has shown, first, that human beings rely very heavily on heuristics and, second, that we often have too much confidence in them. The result is that our probability judgments often go very wrong, and sometimes our thinking gets utterly mixed up. In this regard, two heuristics are particularly instructive: the representativeness heuristic and the availability heuristic.

*THE REPRESENTATIVENESS HEURISTIC.* A simple example illustrates how errors can arise from the representativeness heuristic. Imagine that you are randomly dealt five-card hands from a standard deck. Which of the following two hands is more likely to come up?

HAND #1	HAND #2
Three of clubs	Ace of spades
Seven of diamonds	Ace of hearts
Nine of diamonds	Ace of clubs
Queen of hearts	Ace of diamonds
King of spades	King of spades

A surprisingly large number of people will automatically say that the second hand is much less likely than the first. Actually, if you think about it a little, it should be obvious that any two specific hands have exactly the same likelihood of being dealt in a fair game. Here people get confused because the first hand is unimpressive; and, because unimpressive hands come up all the time, it strikes us as a representative hand. In many card games, however, the second hand is very impressive—something worth talking about—and thus looks unrepresentative. Our reliance on representativeness blinds us to a simple and obvious point about probabilities: Any specific hand is as likely to occur as any other.

**DISCUSSION QUESTION**

Linda is thirty-one years old, single, outspoken, and very bright. As a student, she majored in philosophy, was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations. Rank the following statements with respect to the probability that they are also true of Linda, then explain your rankings:

- Linda is a teacher in elementary school.
- Linda works in a bookstore and takes yoga classes.
- Linda is active in the feminist movement.
- Linda is a psychiatric social worker.
- Linda is a bank teller.
- Linda is an insurance salesperson.
- Linda is a bank teller and is active in the feminist movement.<sup>1</sup>

*THE AVAILABILITY HEURISTIC.* Because sampling and taking surveys is costly, we often do it imaginatively, that is, in our heads. If you ask a baseball fan which team has the better batting average, Detroit or San Diego, that person might just remember, might go look it up, or might think about each team and try to decide which has the most good batters. The latter approach, needless to say, would be a risky business, but many baseball fans have remarkable knowledge of the batting averages of top hitters. Even with this knowledge, however, it is easy to go wrong. The players that naturally come to mind are the stars on each team. They are more available to our memory, and we are likely to make our judgment on the basis of them alone. Yet such a sample can easily be biased because all the batters contribute to the team average, not just the stars. The fact that the weak batters on one team are much better than the weak batters on the other can swing the balance.

**DISCUSSION QUESTION**

In four pages of a novel (about 2,000 words), how many words would you expect to find that have the form       n       (seven-letter words with “n” in the sixth place)? Write down your answer. Now, how many words would you expect to find that have the form       ing (seven-letter words that end with “ing”)? Explain your answers.<sup>2</sup>

The point of examining these heuristics and noting the errors that they produce is not to suggest that we should cease relying on them. First, there is a good chance that this would be psychologically impossible, because the use of such heuristics seems to be built into our psychological makeup. Second, over a wide range of standard cases, these heuristics give quick and largely accurate estimates. Difficulties typically arise in using these

heuristics when the situation is nonstandard—that is, when the situation is complex or out of the ordinary.

To avoid such mistakes when making important judgments about probabilities, we need to ask, “Is the situation sufficiently standard to allow the use of heuristics?” Because this is a mouthful, we might simply ask, “Is this the sort of thing that people can figure out in their heads?” When the answer to that question is “No,” as it often is, then we need to turn to more formal procedures for determining probabilities.

#### DISCUSSION QUESTION

In a remarkable study,<sup>3</sup> Thomas Gilovich, Robert Vallone, and Amos Tversky found a striking instance of people’s tendency to treat things as statistically significant when they are not. In professional basketball, certain players have the reputation of being streak shooters. Streak shooters seem to score points in batches, then go cold and are not able to buy a basket. Stated more precisely, in streak shooting, “the performance of a player during a particular period is significantly better than expected on the basis of the player’s overall record.” To test whether streak shooting really exists, the authors made detailed study of a year’s shooting record for the players on the Philadelphia 76ers. This team included Andrew Toney, noted around the league as being streak shooter. The authors found no evidence for streak shooting, not even for Andrew Toney. How would you go about deciding whether streak shooting exists or not? If, as Gilovich, Vallone, and Tversky have argued, belief in streak shooting is a “cognitive illusion,” why do so many people, including most professional athletes, believe that it does exist?

## THE LANGUAGE OF PROBABILITY

The first step in figuring out probabilities is to adopt a more precise way of talking. Our common language includes various ways of expressing probabilities. Some of the guard terms discussed in Chapter 3 provide examples of informal ways of expressing probability commitments. Thus, someone might say that it is unlikely that the New England Patriots will win the Super Bowl this year without saying how unlikely it is. We can also specify various degrees of probability. Looking out the window, we might say that there is a fifty-fifty chance of rain. More vividly, someone might have remarked that Ron Paul does not have a snowball’s chance in hell of ever winning a presidential election. In each case, the speaker is indicating the relative strength of the evidence for the occurrence or nonoccurrence of some event. To say that there is a fifty-fifty chance that it will rain indicates that we hold that the evidence is equally strong that it will rain rather than not rain. The metaphor in the third statement indicates that the person who uttered it believed that the probability of Ron Paul winning a presidential election is essentially nonexistent.

We can make our probability claims more precise by using numbers. Sometimes we use percentages. For example, the weather bureau might say that there is a 75 percent chance of snow tomorrow. This can naturally be changed to a fraction: The probability is  $\frac{3}{4}$  that it will snow tomorrow. Finally, this fraction can be changed to a decimal expression: There is a 0.75 probability that it will snow tomorrow.

The probability scale has two end points: the absolute certainty that the event will occur and the absolute certainty that it will not occur. Because you cannot do better than absolute certainty, a probability can neither rise above 100 percent nor drop below 0 percent (neither above 1, nor below 0). (This should sound fairly obvious, but it is possible to become confused when combining percentages and fractions, as when Yogi Berra was supposed to have said that success is one-third talent and 75 percent hard work.) Of course, what we normally call probability claims usually fall between these two end points. For this reason, it sounds somewhat peculiar to say that there is a 100 percent chance of rain and just plain weird to say the chance of rain is 1 out of 1. Even so, these peculiar ways of speaking cause no procedural difficulties and rarely come up in practice.

## A PRIORI PROBABILITY

---

When people make probability claims, we have a right to ask why they assign the probability they do. In Chapter 8, we saw how statistical procedures can be used for establishing probability claims. Here we will examine the so-called *a priori* approach to probabilities. A simple example will bring out the differences between these two approaches. We might wonder what the probability is of drawing an ace from a standard deck of fifty-two cards. Using the procedure discussed in Chapter 8, we could make a great many random draws from the deck (replacing the card each time) and then form a statistical generalization concerning the results. We would discover that an ace tends to come up roughly one-thirteenth of the time. From this we could draw the conclusion that the chance of drawing an ace is one in thirteen.

But we do not have to go to all this trouble. We can assume that each of the fifty-two cards has an equal chance of being selected. Given this assumption, an obvious line of reasoning runs as follows: There are four aces in a standard fifty-two-card deck, so the probability of selecting one randomly is four in fifty-two. That reduces to one chance in thirteen. Here the set of favorable outcomes is a subset of the total number of equally likely outcomes, and to compute the probability that the favorable outcome will occur, we merely divide the number of favorable outcomes by the total number of possible outcomes. This fraction gives us the probability that the event will occur on a random draw. Since all outcomes here are equally likely,

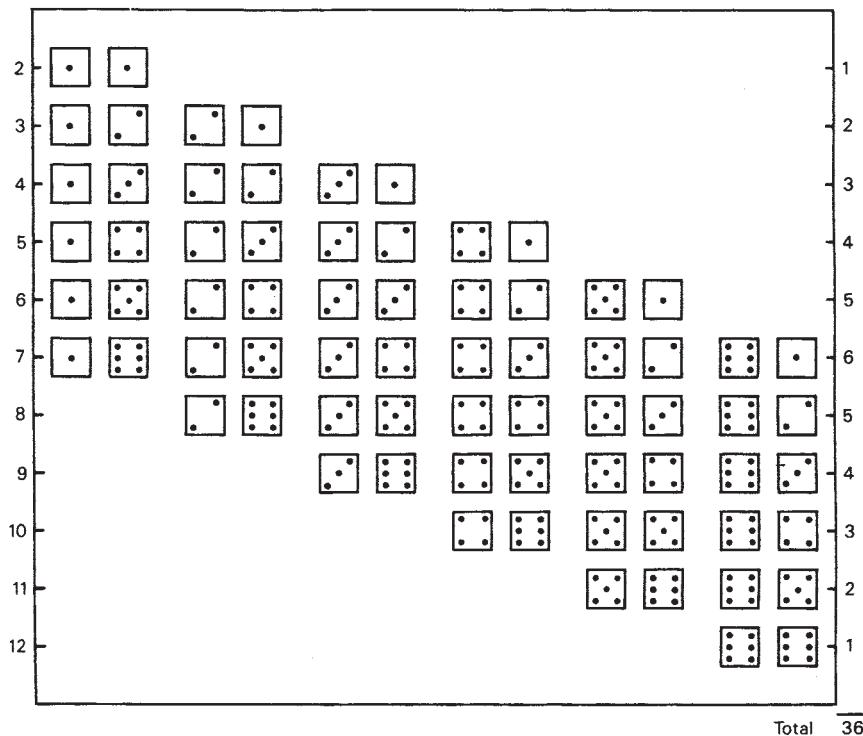
$$\text{Probability of drawing an ace} = \frac{\text{number of aces}}{\text{total number of cards}} = \frac{4}{52} = \frac{1}{13}$$

Notice that in coming to our conclusion that there is one chance in thirteen of randomly drawing an ace from a fifty-two-card deck, we used only mathematical reasoning. This illustrates the a priori approach to probabilities. It is called the a priori approach because we arrive at the result simply by reasoning about the circumstances.

In calculating the probability of drawing an ace from a fifty-two-card deck, we took the ratio of favorable equally likely outcomes to total equally likely outcomes. Generally, then, the probability of a hypothesis  $h$ , symbolized "Pr( $h$ )," when all outcomes are equally likely, is expressed as follows:

$$\text{Pr}(h) = \frac{\text{favorable outcomes}}{\text{total outcomes}}$$

We can illustrate this principle with a slightly more complicated example. What is the probability of throwing an eight on the cast of two dice? The following table lists all of the equally likely ways in which two dice can turn up on a single cast. Notice that five of the thirty-six possible outcomes produce an eight. Hence, the probability of throwing an eight is 5/36.



**EXERCISE I**

Using the above chart, answer the following questions about the total on throw of two dice:

1. What is the probability of throwing a five?
2. Which number has the highest probability of being thrown? What is its probability?
3. What is the probability of throwing an eleven?
4. What is the probability of throwing either a seven or an eleven?
5. Which is more likely: throwing either a five or an eight?
6. Which is more likely: throwing a five or an eight, or throwing a two or a seven?
7. What is the probability of throwing a ten or above?
8. What is the probability of throwing an even number?
9. What is the probability of throwing an odd number?
10. What is the probability of throwing a value from four to six?
11. What is the probability of throwing either a two or a twelve?
12. What is the probability of throwing a value from two to twelve?

**SOME RULES OF PROBABILITY**

Suppose you have determined the probability that certain simple events will occur; how do you go about applying this information to combinations of events? This is a complex question and one that can be touched on only lightly in this text. There are, however, some simple rules of probability that are worth knowing because they can guide us in making choices when outcomes are uncertain.

**PROBABILITIES OF NEGATIONS**

By convention, events are assigned probabilities between 0 and 1 (inclusive). An event is going to either occur or not occur; that, at least, is certain (that is, it has a probability of 1). From this it is easy to see how to calculate the probability that the event will *not* occur, given the probability that it will occur: We simply subtract the probability that it will occur from 1. This is our first rule:

**RULE 1: NEGATION.** The probability that an event will *not* occur is 1 minus the probability that it will occur. Symbolically:

$$\Pr(\text{not } h) = 1 - \Pr(h)$$

For example, the probability of drawing an ace from a standard deck is one in thirteen, so the probability of *not* drawing an ace is twelve in thirteen.

This makes sense because there are forty-eight out of fifty-two ways of not drawing an ace, and this reduces to twelve chances in thirteen.

## PROBABILITIES OF CONJUNCTIONS

We often want to know not just how likely it is that one single event will occur but, instead, how likely it is that two events will occur together in a certain order. Here's a simple rule for calculating probabilities in some such cases:

**RULE 2: CONJUNCTION WITH INDEPENDENCE.** Given two independent events, the probability of their *both* occurring is the product of their individual probabilities. Symbolically (where  $h_1$  and  $h_2$  are independent):

$$\Pr(h_1 \& h_2) = \Pr(h_1) \times \Pr(h_2)$$

Here the word "independent" needs explanation. Suppose you randomly draw a card from the deck, then put it back, shuffle, and draw again. In this case, the outcome of the first draw provides no information about the outcome of the second draw, so it is *independent* of it. What is the probability of drawing two aces in a row using this system? Using Rule 2, we see that the answer is  $1/13 \times 1/13$ , or 1 chance in 169.

The situation is different if we do *not* replace the card after the first draw. Rule 2 does not apply to this case because the two events are no longer independent. The chances of getting an ace on the first draw are still one in thirteen, but if an ace is drawn and not returned to the pack, then there is one less ace in the deck, so the chances of drawing an ace on the next draw are reduced to three in fifty-one. Thus, the probability of drawing two consecutive aces without returning the first draw to the deck is  $4/52 \times 3/51$ , or 1 in 221, which is considerably lower than 1 in 169.

If we want to extend Rule 2 to cover cases in which the events are *not* independent, then we will have to speak of the probability of one event occurring, given that another has occurred. The probability that  $h_2$  will occur given that  $h_1$  has occurred is called the *conditional probability* of  $h_2$  on  $h_1$  and is usually symbolized thus:  $\Pr(h_2 | h_1)$ . This probability is calculated by considering only those cases where  $h_1$  is true and then dividing the number of cases within that group where  $h_2$  is also true by the total number of cases in that group. Symbolically:

$$\Pr(h_2 | h_1) = \frac{\text{favorable outcomes where } h_1}{\text{total outcomes where } h_1} = \frac{\text{outcomes where } h_1 \text{ and } h_2}{\text{total outcomes where } h_1}$$

Using this notion of conditional probability, Rule 2 can be modified as follows to deal with cases in which events need not be independent:

**RULE 2G: CONJUNCTION IN GENERAL.** Given two events, the probability of their both occurring is the probability of the first occurring times the probability of the second occurring, given that the first has occurred. Symbolically:

$$\Pr(h_1 \& h_2) = \Pr(h_1) \times \Pr(h_2 | h_1)$$

Notice that, in the event that  $h_1$  and  $h_2$  are independent, the probability of  $h_2$  is not related to the occurrence of  $h_1$ , so the probability of  $h_2$  on  $h_1$  is simply the probability of  $h_2$ . Thus, Rule 2 is simply a special case of the more general Rule 2G.

We can extend these rules to cover more than two events. For example, with Rule 2, regardless of the number of events we might consider, provided that they are independent of each other, the probability of all of them occurring is the product of each one of them occurring. For example, the chances of flipping a coin and having it come up heads is one chance in two. What are the chances of flipping a coin eight times and having it come up heads every time? The answer is:

$$1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2$$

which equals 1 chance in 256.

## PROBABILITIES OF DISJUNCTIONS

Our next rule allows us to answer questions of the following kind: What are the chances of *either* an eight or a two coming up on a single throw of the dice? Going back to the chart, we saw that we could answer this question by counting the number of ways in which a two can come up (which is one) and adding this to the number of ways in which an eight can come up (which is five). We could then conclude that the chances of one or the other of them coming up are six in thirty-six, or 1/6. The principle involved in this calculation can be stated as follows:

**RULE 3: DISJUNCTION WITH EXCLUSIVITY.** The probability that at least one of two mutually exclusive events will occur is the sum of the probabilities that each of them will occur. Symbolically (where  $h_1$  and  $h_2$  are mutually exclusive):

$$\Pr(h_1 \text{ or } h_2) = \Pr(h_1) + \Pr(h_2)$$

To say that events are *mutually exclusive* means that they cannot both occur together. You cannot, for example, get both a two and an eight on the same cast of two dice. You might, however, throw neither a two nor an eight, since you might throw some other number.

When events are not mutually exclusive, the rule for calculating disjunctive probabilities becomes more complicated. Suppose, for example, that exactly

half the class is female and exactly half the class is over nineteen and the age distribution is the same for females and males. What is the probability that a randomly selected student will be either a female or over nineteen? If we simply add the probabilities ( $1/2 + 1/2 = 1$ ), we would get the result that we are certain to pick someone who is either female or over nineteen. But that answer is wrong, because a quarter of the class is male and not over nineteen, and one of them might have been randomly selected. The correct answer is that the chances are  $3/4$  of randomly selecting someone who is either female or over nineteen.

We can see that this is the correct answer by examining the following table:

	Over Nineteen	Not over Nineteen
Female	25%	25%
Male	25%	25%

It is easy to see that in 75 percent of the cases, a randomly selected student will be either female or over nineteen. The table also shows what went wrong with our initial calculation. The top row shows that 50 percent of the students are female. The left column shows that 50 percent of the students are over nineteen. But we cannot simply add these figures to get the probability of a randomly selected student being either female or over nineteen. Why? Because that would double-count the females over nineteen. We would count them once in the top row and then again in the left column. To compensate for such double-counting, we need to subtract the students who are both female and over nineteen. The upper left figure shows that this is 25%. So the correct way to calculate the answer is  $50\% + 50\% - 25\% = 75\%$ .

This pattern is reflected in the general rule governing the calculation of disjunctive probabilities:

**RULE 3G: DISJUNCTION IN GENERAL.** The probability that at least one of two events will occur is the sum of the probabilities that each of them will occur, minus the probability that they will both occur. Symbolically:

$$\Pr(h_1 \text{ or } h_2) = \Pr(h_1) + \Pr(h_2) - \Pr(h_1 \& h_2)$$

If  $h_1$  and  $h_2$  are mutually exclusive, then  $\Pr(h_1 \& h_2) = 0$ , and Rule 3G reduces to Rule 3. Thus, as with Rules 2 and 2G, Rule 3 is simply a special case of the more general Rule 3G.

## PROBABILITIES IN A SERIES

Before stating Rule 4, we can think about a particular example. What is the probability of tossing heads at least once in eight tosses of a coin? Here it is tempting to reason in the following way: There is a 50 percent chance of getting heads on the first toss and a 50 percent chance of getting heads on the second toss, so after two tosses it is already certain that we will toss

heads at least once, and thus after eight tosses there should be a 400 percent chance. In other words, you cannot miss. There are two good reasons for thinking that this argument is fishy. First, probability can never exceed 100 percent. Second, there must be some chance, however small, that we could toss a coin eight times and not have it come up heads.

The best way to look at this question is to restate it so that the first two rules can be used. Instead of asking what the probability is that heads will come up at least once, we can ask what the probability is that heads will *not* come up at least once. To say that heads will not come up even once is equivalent to saying that tails will come up eight times in a row. By Rule 2, we know how to compute that probability: It is  $1/2$  multiplied by itself eight times, and that, as we saw, is  $1/256$ . Finally, by Rule 1 we know that the probability that this will not happen (that heads will come up at least once) is  $1 - (1/256)$ . In other words, the probability of tossing heads at least once in eight tosses is  $255/256$ . That comes close to a certainty, but it is not quite a certainty.

We can generalize these results as follows:

**Rule 4: Series with Independence.** The probability that an event will occur at least once in a series of independent trials is 1 minus the probability that it will not occur in that number of trials. Symbolically (where  $n$  is the number of independent trials):

$$\Pr(h \text{ at least once in } n \text{ trials}) = 1 - \Pr(\text{not } h)^n$$

Strictly speaking, Rule 4 is unnecessary, since it can be derived from Rules 1 and 2, but it is important to know because it blocks a common misunderstanding about probabilities: People often think that something is a sure thing when it is not.

## PERMUTATIONS AND COMBINATIONS

Another common confusion is between permutations and combinations. A *permutation* is a set of items whose order is specified. A *combination* is a set of items whose order is not specified. Imagine, for example, that three cards—the jack, queen, and king of spades—are facedown in front of you. If you pick two of these cards in turn, there are three possible combinations: jack and queen, jack and king, and queen and king. In contrast, there are six possible permutations: jack then queen, queen then jack, jack then king, king then jack, queen then king, and king then queen.

Rule 2 is used to calculate probabilities of permutations—of conjunctions of events in a particular order. For example, if you flip a fair coin twice, what is the probability of its coming up heads and tails in that order (that is, heads on the first flip and tails on the second flip)? Since the flips are independent, Rule 2 tells us that the answer is  $1/2 \times 1/2 = 1/4$ . This answer is easily confirmed by counting the possible permutations (heads then heads, heads then tails, tails then heads, tails then tails). Only one of these four permutations (heads then tails) is a favorable outcome.

We need to calculate probabilities of combinations in a different way. For example, if you flip a fair coin twice, what is the probability of its landing heads and tails in *any* order? There are two ways for this to happen. The coin could come up either heads then tails or tails then heads. These alternatives are mutually exclusive, so the probability of this disjunction by Rule 3 is  $1/4 + 1/4 = 1/2$ . This is confirmed by counting two possibilities (heads then tails, tails then heads) out of four (heads then heads, heads then tails, tails then heads, tails then tails). Another way to calculate this probability is to realize that the first flip doesn't matter. Whatever you get on the first flip (heads or tails), you need the opposite on the second flip. You are certain to get either heads or tails on the first flip, so this probability is 1. Then, regardless of what happens on the first flip, the probability of getting the opposite on the second flip is  $1/2$ . These results are independent, so the probability of their conjunction by Rule 2 is the product  $1 \times 1/2 = 1/2$ .

We can also use our rules to calculate probabilities of combinations without independence. Rule 2G tells us that the probability of drawing an ace, not putting this card back in the deck, and then drawing a king is  $4/52 \times 4/51 = 16/2,652$ . But what is the probability of drawing an ace and a king in any order? It is the probability of drawing either an ace or a king and then drawing the other one, given that you drew the first one. That probability by Rule 2G is  $8/52 \times 4/51 = 32/2,652$ . The difference between this result and the previous one, where the order was specified, shows why we need to determine whether we are dealing with permutations or combinations.

### EXERCISE II

Use the rules of probability to calculate these probabilities:

1. What is the probability of rolling a five on one throw of a fair six-sided die?
2. What is the probability of *not* rolling a five on one throw of a fair six-sided die?
3. If you roll a five on your first throw of a fair six-sided die, what is probability of rolling *another* five on a second throw of that die?
4. If you roll two fair six-sided dice one time, what are the chances that *both* of the dice will come up a five?
5. If you roll two fair six-sided dice one time, what are the chances that *one or the other* (or both) of the dice will come up a five?
6. If you roll two fair six-sided dice one time, what are the chances that *one and only one* of the dice will come up a five?
7. If you roll two fair six-sided dice one time, what are the chances that *at least one* of the dice will come up a five?
8. If you roll two fair six-sided dice one time, what are the chances that *at least one* of the dice will *not* come up a five?
9. If you roll six fair six-sided dice one time, what are the chances that *at least one* of the dice will come up a five?
10. If you roll six fair six-sided dice one time, what are the chances that *at least one* of the dice will *not* come up a five?

**EXERCISE III**

Compute the probability of making the following draws from a standard fifty-two-card deck:

1. Drawing either a seven or a five on a single draw.
2. Drawing neither a seven nor a five on a single draw.
3. Drawing a seven and then, without returning the first card to the deck, drawing a five on the next draw.
4. Same as 3, but the first card is returned to the deck and the deck is shuffled after the first draw.
5. Drawing at least one spade in a series of three consecutive draws, when the card drawn is not returned to the deck.
6. Drawing at least one spade in a series of four consecutive draws, when the card drawn is not returned to the deck.
7. Same as 6, but the card is returned to the deck after each draw and the deck is reshuffled.
8. Drawing a heart and a diamond in that order in two consecutive draws, when the first card is returned to the deck and the deck is reshuffled the first draw.
9. Drawing a heart and a diamond in any order in two consecutive draws, when the first card is returned to the deck and the deck is reshuffled the first draw.
10. Drawing a heart and a diamond in any order in two consecutive draws, when the first card is not returned to the deck after the first draw.

**EXERCISE IV**

Suppose there are two little lotteries in town, each of which sells exactly one hundred tickets.

1. If each lottery has only one winning ticket, and you buy two tickets to the *same* lottery, what is the probability that you will have a winning ticket?
2. If each lottery has only one winning ticket, and you buy one ticket to *each* of the two lotteries, what is the probability that you will have at least one winning ticket?
3. If each lottery has only one winning ticket, and you buy one ticket to each of the two lotteries, what is the probability that you will have *two* winning tickets?
4. If each lottery has two winning tickets, and you buy one ticket to each of the two lotteries, what is the probability that you will have at least one winning ticket?
5. If each lottery has two winning tickets, and you buy two tickets to the same lottery, what is the probability that you will have two winning tickets?
6. If each lottery has two winning tickets, and you buy two tickets to the same lottery, what is the probability that you will have at least one winning ticket?

**EXERCISE V**

1. You are presented with two bags, one containing two ham sandwiches and the other containing a ham sandwich and a cheese sandwich. You reach in one bag and draw out a ham sandwich. What is the probability that the other sandwich in the bag is also a ham sandwich?
2. You are presented with three bags: two contain a chicken-fat sandwich and one contains a cheese sandwich. You are asked to guess which bag contains the cheese sandwich. You do so, and the bag you selected is set aside. (You obviously have one chance in three of guessing correctly.) From the two remaining bags, one containing a chicken-fat sandwich is then removed. You are now given the opportunity to switch your selection to the remaining bag. Will such a switch increase, decrease, or leave unaffected your chances of correctly ending up with the bag with the cheese sandwich in it?

**BAYES'S THEOREM**

Although dice and cards provide nice, simple models for learning how to calculate probabilities, real life is usually more complicated. One particularly interesting and important form of problem arises often in medicine. Suppose that Wendy tests positive for colon cancer. The treatment for colon cancer is painful and dangerous, so, before subjecting Wendy to that treatment, her doctor wants to determine how likely it is that Wendy really has colon cancer. After all, no test is perfect. Regarding the test that was used on Wendy, previous studies have revealed the following probabilities:

The probability that a person in the general population has colon cancer is 0.3 percent (or 0.003).

If a person has colon cancer, then the probability that the test is positive is 90 percent (or 0.9).

If a person does not have colon cancer, then the probability that the test is positive is 3 percent (or 0.03).

On these assumptions, what is the probability that Wendy actually has colon cancer, given that she tested positive? Most people guess that this probability is fairly high. Even most trained physicians would say that Wendy probably has colon cancer.<sup>4</sup>

What is the correct answer? To calculate the probability that a person who tests positive actually has colon cancer, we need to divide the number of favorable outcomes by the number of total outcomes. The favorable outcomes include everyone who tests positive and really has colon cancer. This outcome is not “favorable” to Wendy, so we will describe this group as *true positives*. The total outcomes include everyone who tests positive. This

includes the true positives plus the *false positives*, which are those who test positive but do not have colon cancer. Given the stipulated probabilities, in a normal population of 100,000 people, there will be 270 true positives ( $100,000 \times 0.003 \times 0.9$ ) and 2,991 false positives [ $(100,000 - 300) \times 0.03$ ]. Thus, the probability that Wendy has colon cancer is about  $270/(270 + 2,991)$ . That is only about 8.3 percent, when most people estimate above 50 percent!

Why do people, including doctors, overestimate these probabilities so badly? Part of the answer seems to be that they focus on the rate of true positives (90 percent) and forget that, because there are so many people without colon cancer (99.7 percent of the total population), even a small rate of false positives (3 percent) will yield a large number of false positives (2,991) that swamps the much smaller number of true positives (270). (When the question about probability was reformulated in terms of the number of people in each group, most doctors came up with the correct answer.) For whatever reason, people have a strong tendency to make mistakes in cases like these, so we need to be careful, especially when so much is at stake.

One way to calculate probabilities like these uses a famous theorem that was first presented by an English clergyman named Thomas Bayes (1702–1761). A simple proof of this theorem applies the laws of probability from the preceding section. We want to figure out  $\Pr(h | e)$ , that is, the probability of the hypothesis  $h$  (for example, Wendy has colon cancer), given the evidence  $e$  (for example, Wendy tested positive for colon cancer). To get there, we start from Rule 2G:

$$1. \Pr(e \& h) = \Pr(e) \times \Pr(h | e)$$

Dividing both sides by  $\Pr(e)$  gives us:

$$2. \Pr(h|e) = \frac{\Pr(e \& h)}{\Pr(e)}$$

If two formulas are logically equivalent, they must have the same probability. We can establish by truth tables (as in Chapter 6) that “ $e$ ” is logically equivalent to “ $(e \& h) \vee (e \& \sim h)$ .” Consequently, we may replace “ $e$ ” in the denominator of item 2 with “ $(e \& h) \vee (e \& \sim h)$ ” to get:

$$3. \Pr(h|e) = \frac{\Pr(e \& h)}{\Pr[(e \& h) \vee (e \& \sim h)]}$$

Since “ $e \& h$ ” and “ $e \& \sim h$ ” are mutually exclusive, we can apply Rule 3 to the denominator of item 3 to get:

$$4. \Pr(h|e) = \frac{\Pr(e \& h)}{\Pr(e \& h) + \Pr(e \& \sim h)}$$

Finally, we apply Rule 2G to item 4 and get:

$$\text{BT: } \Pr(h | e) = \frac{\Pr(h) \times \Pr(e | h)}{[\Pr(h) \times \Pr(e | h)] + \Pr(\sim h) \times \Pr(e | \sim h)}$$

This is a simplified version of Bayes's theorem.

This theorem enables us to calculate the desired probability in our original example:

$h$  = the patient has colon cancer

$e$  = the patient tests positive for colon cancer

$\Pr(h) = 0.003$

$\Pr(\sim h) = 1 - \Pr(h) = 0.997$

$\Pr(e | h) = 0.9$

$\Pr(e | \sim h) = 0.03$

If we substitute these values into Bayes's theorem, we get:

$$\Pr(h | e) = \frac{0.003 \times 0.9}{[0.003 \times 0.9] + [0.997 \times 0.03]} = \text{about } 0.083$$

In this way, we can calculate the conditional probability of the hypothesis given the evidence from its reverse, that is, from the conditional probability of the evidence given the hypothesis. That is what makes Bayes's theorem so useful.

Many people find a different method more intuitive. The first step is to set up a table. The two factors to be related are: (1) whether the patient has colon cancer and (2) whether the patient tests positive for colon cancer. To chart all possible combinations of these two factors, we need a table like this:

	Colon Cancer	Not Colon Cancer	Total
Test Positive			
Do Not Test Positive			
Total			

Next, we need to enter a population size in the lower right box. The probabilities will not be affected by the population size, but it is cleaner to pick a population that is large enough to get whole numbers when the population is multiplied by the given probabilities. To determine the right size population, add the number of places to the right of the decimal point in the two most specific probabilities, then pick a population of 10 to the power of that sum. In our example, the most specific probabilities are 0.003 and 0.03, and  $3 + 2 = 5$ , so we can enter  $10^5$ :

	<b>Colon Cancer</b>	<b>Not Colon Cancer</b>	<b>Total</b>
<b>Test Positive</b>			
<b>Do Not Test Positive</b>			
<b>Total</b>			100,000

This population size represents the total number of people who are tested. We have no information about the ones who are not tested, so they cannot figure into our calculations.

The bottom row can now be filled in by dividing the total population into those who have colon cancer and those who do not have colon cancer. Just multiply the population size by the probability of colon cancer in the general population [ $\Pr(h)$ ] to get a number for the second box on the bottom row. This figure represents the total number of people with colon cancer in this population. Then subtract that product from the population size and put the remainder in the remaining box. This represents the total number of people without colon cancer in this population. Since these two groups exhaust the population, they must add up to the total population size. In our case, we were given that the probability that a person in the general population has colon cancer is 0.003. On this basis, we can fill in the bottom row of the table:

	<b>Colon Cancer</b>	<b>Not Colon Cancer</b>	<b>Total</b>
<b>Test Positive</b>			
<b>Do Not Test Positive</b>			
<b>Total</b>	300	99,700	100,000

Next, fill out the second column by dividing the total number of people with colon cancer into those who test positive and those who do not test positive. These numbers can be calculated with the given conditional probability of testing positive, given colon cancer [ $\Pr(e|h)$ ]. In our example, if a person has colon cancer, the probability that the test is positive is 0.9. Thus,  $270 (= 0.9 \times 300)$  of the people in the colon cancer column will test positive and the rest ( $300 - 270 = 30$ ) will not, so we get these figures:

	<b>Colon Cancer</b>	<b>Not Colon Cancer</b>	<b>Total</b>
<b>Test Positive</b>	270		
<b>Do Not Test Positive</b>	30		
<b>Total</b>	300	99,700	100,000

Similarly, we can fill out the third column by dividing the total number of people without colon cancer into those who test positive and those who do not test positive. Here we use the conditional probability of a positive test, given that a person does not have colon cancer [ $\text{Pr}(e|\sim h)$ ]. This probability was given as 0.03, and  $0.03 \times 99,700 = 2,991$ . This number means that, out of a normal population of 99,700 without colon cancer, 2,991 will test positive. Since the figures in this column must add up to a total of 99,700, the remaining figure is  $99,700 - 2,991 = 96,709$ :

	<b>Colon Cancer</b>	<b>Not Colon Cancer</b>	<b>Total</b>
<b>Test Positive</b>	270	2,991	
<b>Do Not Test Positive</b>	30	96,709	
<b>Total</b>	300	99,700	100,000

Finally, we can fill out the fourth column by calculating total numbers of people who test positive or do not test positive. Simply add across the rows:

	<b>Colon Cancer</b>	<b>Not Colon Cancer</b>	<b>Total</b>
<b>Test Positive</b>	270	2,991	3,261
<b>Do Not Test Positive</b>	30	96,709	96,739
<b>Total</b>	300	99,700	100,000

Check your calculations by adding the right column:  $3,261 + 96,739 = 100,000$ .

Now that our population is divided up, the solution is staring you in the face. This table shows us that, in a normal population of 100,000 tested people distributed according to the given probabilities, a total of 3,261 will test positive. Out of those, 270 will have colon cancer. Thus, the probability that the patient has colon cancer, given that this patient tested positive, is  $270/3,261$ , which is about 0.083 or 8.3 percent, just as before.

You can also read off other conditional probabilities. If you want to know the conditional probability of *not* having colon cancer, given that your test did *not* come out positive, then you need to look at the row for those who do *not* test positive. The figure at the right of this row tells you that a total of 96,739 out of the total population do not test positive. The column under "Not Colon Cancer" then tells you that 96,709 of these do not have colon cancer. Thus, the conditional probability of not having colon cancer given your test did not come out positive is  $96,709/96,739$  or about 0.9997. That means that, if you test negative, the odds are extremely high that you do not have colon cancer.

Tables like these work by dividing the population into groups. We already learned some names for these groups:

	Hypothesis ( $h$ )	Not Hypothesis ( $\sim h$ )	
Evidence ( $e$ )	True Positives	False positives	
Not Evidence ( $\sim e$ )	False Negatives	True Negatives	
			Population

False positives are sometimes also called *false alarms*, and false negatives are sometimes called *misses*. A little more terminology is also common:

$\Pr(h)$  = base rate or prevalence or prior probability

$\Pr(h | e)$  = solution or posterior probability

$\Pr(e | h)$  = sensitivity of the test

$\Pr(\sim e | \sim h)$  = specificity of the test

$1 - \Pr(e | h) = 1 - \text{sensitivity} = \text{false negative rate}$

$1 - \Pr(\sim e | \sim h) = 1 - \text{specificity} = \text{false positive rate}$

You don't need to use these terms in order to calculate the probabilities, but it is useful to learn them so that you will be able to understand people who discuss these issues.

One of the most important lessons of Bayes's theorem is that the base rate has big effects. To see how much it matters, let's recalculate the solution [ $\Pr(h | e)$ ] in our colon cancer example for different values of the base rate [ $\Pr(h)$ ] using the same test with the same sensitivity ( $\Pr(e | h) = 0.9$ ) and specificity [ $\Pr(\sim e | \sim h) = 0.97$ ]:

If  $\Pr(h) = 0.003$ , then  $\Pr(h | e) = 0.083$

If  $\Pr(h) = 0.03$ , then  $\Pr(h | e) = 0.48$

If  $\Pr(h) = 0.3$ , then  $\Pr(h | e) = 0.93$

#### EXERCISE VI

Construct tables to confirm these calculations of  $\Pr(h | e)$  for base rates of 0.03 and 0.3.

These calculations show that a positive test result for a given test means a lot more when the base rate is high than when it is low. Thus, if doctors use the specified test as a *screening test* in the general population, and if the rate of colon cancer in that general population is only 0.003, then a positive test result by itself does not show that the patient has cancer. In contrast, if doctors use the specified test as a *diagnostic test* only for people with certain symptoms, and if the rate of colon cancer among people with

those symptoms is 0.3, then a positive test result does show that the patient probably has cancer, though the test still might be mistaken. Bayes's theorem, thus, reveals the right ways and the wrong ways to use and interpret such tests.

Notice also what happens to the probabilities when additional tests are performed. In our original example, one positive test result raises the probability of cancer from the base rate of 0.003 to our solution of 0.083. Now suppose that the doctor orders an additional independent test, and the result is again positive. To apply Bayes's theorem at this point, we can take the probability after the original positive test result (0.083) as the prior probability or base rate in calculating the probability after the second positive test result. This method makes sense because we are now interested not in the general population but only in the subpopulation that already tested positive on the first test. The solution after two tests [ $\text{Pr}(h | e)$ ], where “*e*” is now two independent positive test results in a row, is 0.731. Next, if the doctor orders a third independent test, and if the result is positive yet again, then  $\text{Pr}(h | e)$  increases to 0.988. Bayes's theorem, thus, reveals the technical rationale behind the commonsense practice of ordering additional tests. Problems arise only when doctors put too much faith in a single positive test result without doing any additional tests.

#### EXERCISE VII

Construct tables to confirm the above calculations of probabilities after a second and third positive test result.

#### EXERCISE VIII

1. What would Wendy's chances of having colon cancer be if the other probabilities remained the same as in the original example, except that the probability that a person in the general population has colon cancer only 0.1 percent (or 0.001)?
2. What would Wendy's chances of having colon cancer be if the other probabilities remained the same as in the original example, except that the probability that a person in the general population has colon cancer 1 percent (0.01)?
3. What would Wendy's chances of having colon cancer be if the other probabilities remained the same as in the original example, except that the conditional probability that the test is positive, given that the patient has colon cancer, is only 50 percent (or 0.5)?
4. What would Wendy's chances of having colon cancer be if the other probabilities remained the same as in the original example, except that the conditional probability that the test is positive, given that the patient has colon cancer, is 99 percent (or 0.99)?

(continued)

5. What would Wendy's chances of having colon cancer be if the other probabilities remained the same as in the original example, except that the conditional probability that the test is positive, given that the patient does not have colon cancer, is 1 percent (or 0.01)?
6. What would Wendy's chances of having colon cancer be if the other probabilities remained the same as in the original example, except that the conditional probability that the test is positive, given that the patient does not have colon cancer, is 10 percent (0.1)?
7. Chris tested positive for cocaine once in a random screening test. This test has a sensitivity and specificity of 95 percent, and 20 percent of the students in Chris's school use cocaine. What is the probability that Chris really did use cocaine?
8. As in problem 7, 20 percent of the students in Chris's school use cocaine, but this time Chris tests positive for cocaine on two independent tests, both of which have a sensitivity and specificity of 95 percent. Now what is the probability that Chris really did use cocaine?
9. In your neighborhood, 20 percent of the houses have high levels of radon gas in their basements, so you ask an expert to test your basement. An inexpensive test comes out positive in 80 percent of the basements that actually have high levels of radon, but it also comes out positive in 10 percent of the basements that do not have high levels of radon. If this inexpensive test comes out positive in your basement, what is the probability that there is a high level of radon gas in your basement?
10. A more expensive test for radon is also more accurate. It comes out positive in 99 percent of the basements that actually have high levels of radon. It also tests positive in 2 percent of the basements that do not have high levels of radon. As in problem 7, 20 percent of the houses in your neighborhood have radon in their basement. If the expensive test comes out positive in your basement, what is the probability that there is a high level of radon gas in your basement?
11. Late last night a car ran into your neighbor and drove away. In your town, there are 500 cars, and 2 percent of them are Porsches. The only eyewitness to the hit-and-run says the car that hit your neighbor was a Porsche. Tested under similar conditions, the eyewitness mistakenly classifies cars of other makes as Porsches 10 percent of the time, and correctly classifies Porsches as such 80 percent of the time. What are the chances that the car that hit your neighbor really was a Porsche?
12. Late last night a dog bit your neighbor. In your town, there are 400 dogs, 95 percent of them are black Labrador retrievers, and the rest are pit bulls. The only eyewitness to the event, a veteran dog breeder, says that the dog who bit your neighbor was a pit bull. Tested under similar low-light conditions, the eyewitness mistakenly classifies black Labs as pit bulls only 2 percent of the time, and correctly classifies pit bulls as pit bulls 90 percent of the time. What are the chances that dog who bit your neighbor really was a pit bull?

13. In a certain school, the probability that a student reads the assigned pages before a lecture is 80 percent (or 0.8). If a student does the assigned reading in advance, then the probability that the student will understand the lecture is 90 percent (or 0.9). If a student does not do the assigned reading in advance, then the probability that the student will understand the lecture is 10 percent (or 0.1). What is the probability that a student did the reading in advance, given that she did understand the lecture? What is the probability that a student did *not* do the reading in advance, given that she did *not* understand the lecture?
14. In a different school, the probability that a student reads the assigned pages before a lecture is 60 percent (or 0.6). If a student does the assigned reading in advance, then the probability that, when asked, the student will tell the professor that he did the reading is 100 percent (or 1.0). If a student does not do the assigned reading in advance, then the probability that, when asked, the student will tell the professor that he did the reading is 70 percent (or 0.7). What is the probability that a student did the reading in advance, given that, when asked, he told the professor that he did the reading? What is the probability that a student did not do the reading in advance, given that, when asked, he told the professor that he did not do the reading?

## NOTES

<sup>1</sup> Amos Tversky and Daniel Kahneman, "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *Psychological Review* 90 (1983): 297.

<sup>2</sup> Ibid.

<sup>3</sup> Thomas Gilovich, Robert Vallone, and Amos Tversky, "The Hot Hand in Basketball: The Misperception of Random Sequences," *Cognitive Psychology* 17 (1985): 295–314. The quotation is from pages 295 to 296.

<sup>4</sup> See Gerd Gigerenzer, *Calculated Risk: How to Know When Numbers Deceive You* (New York: Simon & Schuster, 2003).



## CHOICES

*Probabilities are used not only when we determine what to believe but also when we choose what to do. Although we sometimes assume that we know how our actions will turn out, we often have to make decisions in the face of risk, when we do not know what the outcomes of our options will be, but we do know the probabilities of those outcomes. To help us assess reasoning about choices involving risk, this chapter will explain the notions of expected monetary value and expected overall value. Our most difficult choices arise, however, when we do not know even the probabilities of various outcomes. Such decisions under ignorance or uncertainty pose special problems, for which a number of rules have been proposed. Although these rules are useful in many situations, their limitations will also be noted.*

### **EXPECTED MONETARY VALUE**

---

It is obvious that having some sense of probable outcomes is important for running our lives. If we hear that there is a 95 percent chance of rain, this usually provides a good enough reason to call off a picnic. But the exact relationship between probabilities and decisions is complex and often misunderstood.

The best way to illustrate these misunderstandings is to look at lotteries in which the numbers are fixed and clear. A \$1 bet in a lottery might make you as much as \$10 million. That sounds good. Why not take a shot at \$10 million for only a dollar? Of course, there is not much chance of winning the lottery—say, only 1 chance in 20 million—and that sounds bad. Why throw \$1 away on nothing? So we are pulled in two directions. What we want to know is just how good the bet is. Is it, for example, better or worse than a wager in some other lottery? To answer questions of this kind, we need to introduce the notion of expected monetary value.

The idea of expected monetary value takes into account three features that determine whether a bet is financially good or bad: the probability of winning, the net amount you gain if you win, and the net amount you lose if you lose. Suppose that on a \$1 ticket there is 1 chance in 20 million of winning the New York State Lottery, and you will get \$10 million from the state if you do. First, it is worth noticing that, if the state pays you \$10 million,

your net gain on your \$1 ticket is only \$9,999,999. The state, after all, still has your original \$1. So the net gain equals the payoff minus the cost of betting. This is not something that those who win huge lotteries worry about, but taking into account the cost of betting becomes important when this cost becomes high relative to the size of the payoff. There is nothing complicated about the net amount that you lose when you lose on a \$1 ticket: It is \$1.<sup>1</sup>

We can now compute the expected monetary value or financial worth of a bet in the following way:

$$\text{Expected monetary value} =$$

(the probability of winning times the net gain in money of winning) minus  
(the probability of losing times the net loss in money of losing)

In our example, a person who buys a \$1 ticket in the lottery has 1 chance in 20 million of a net gain of \$9,999,999 and 19,999,999 chances in 20 million of a net loss of a dollar. So the expected monetary value of this wager equals:

$$(1/20,000,000 \times \$9,999,999) - (19,999,999/20,000,000 \times \$1)$$

That comes out to  $-\$0.50$ .

What does this mean? One way of looking at it is as follows: If you could somehow buy up all the lottery tickets and thus ensure that you would win, your \$20 million investment would net you \$10 million, or \$0.50 on the dollar—certainly a bad investment. Another way of looking at the situation is as follows: If you invested a great deal of money in the lottery over many millions of years, you could expect to win eventually, but, in the long run, you would be losing fifty cents on every ticket you bought. One last way of looking at the situation is this: You go down to your local drugstore and buy a blank lottery ticket for \$0.50. Since it is blank, you have no chance of winning, with the result that you lose \$0.50 every time you bet. Although almost no one looks at the matter in this way, this is, in effect, what you are doing financially over the long run when you buy lottery tickets.

We are now in a position to distinguish favorable and unfavorable expected monetary values. The expected monetary value is favorable when it is greater than zero. Changing our example, suppose the chances of hitting a \$20 million payoff on a \$1 bet are 1 in 10 million. In this case, the state still has the \$1 you paid for the ticket, so your gain is actually \$19,999,999. The expected monetary value is calculated as follows:

$$(1/10,000,000 \times \$19,999,999) - (9,999,999/10,000,000 \times \$1)$$

That comes to \$1. Financially, this is a good bet, for in the long run you will gain \$1 for every \$1 you bet in such a lottery.

The rule, then, has three parts: (1) If the expected monetary value of the bet is more than zero, then the expected monetary value is *favorable*. (2) If the expected monetary value of the bet is less than zero, then the expected monetary value is *unfavorable*. (3) If the expected monetary value of the bet is zero, then the bet is *neutral*—a waste of time as far as money is concerned.

**EXERCISE I**

Compute the probability and the expected monetary value for the following bets. Each time, you lay down \$1 to bet that a certain kind of card will appear from a standard fifty-two-card deck. If you win, you collect the amount indicated, so your net gain is \$1 less. If you lose, of course, you lose your \$1.

Example: Draw a seven of spades. Win: \$26.

Probability of winning = 1/52

Expected value:  $[1/52 \times (\$26 - \$1)] - (51/52 \times \$1) = -\$0.50$

1. Draw a seven of spades or a seven of clubs. Win: \$26.
2. Draw a seven of any suit. Win: \$26.
3. Draw a face card (jack, queen, or king). Win: \$4.
4. Do *not* draw a face card (jack, queen, or king). Win: \$2.
5. On two consecutive draws without returning the first card to the deck, draw a seven of spades and then a seven of clubs. Win: \$1,989.
6. Same as in problem 5, but this time the card is returned to the deck and the deck is shuffled before the second draw. Win: \$1,989.
7. On two consecutive draws without returning the first card to the deck, do not draw a club. Win: \$1.78.
8. Same as in problem 7, but this time the card is returned to the deck and the deck is shuffled before the second draw. Win: \$1.78.
9. On four consecutive draws without returning any cards to the deck, a seven of spades, then a seven of clubs, then a seven of hearts, and then seven of diamonds. Win: \$1,000,001.
10. On four consecutive draws without returning any cards to the deck, draw four sevens in any order. Win: \$1,000,001.

**EXERCISE II**

Fogelin's Palace in Border, Nevada, offers the following unusual bet. If you win, you make a 50 percent profit on your bet; if you lose, you take a 40 percent loss. That is, if you bet \$1 and win, then you get back \$1.50; if you bet \$1 and lose, you get back \$0.60. The chances of winning are fifty-fifty. This sounds like a marvelous opportunity, but there is one hitch: To play, you must let your bet ride with its winnings, or losses, for four plays. For example, with \$100, a four-bet sequence might look like this:

	<b>Win</b>	<b>Win</b>	<b>Lose</b>	<b>Win</b>
<b>Total</b>	\$150	\$225	\$135	\$202.50

At the end of this sequence, you can pick up \$202.50, and thus make a \$102.50 profit. It seems that Fogelin's Palace is a good place to gamble, but consider

(continued)

the following argument on the other side. Because the chances of winning are fifty-fifty, you will, on the average, win half the time. But notice what happens in such a case:

	<b>Win</b>	<b>Lose</b>	<b>Lose</b>	<b>Win</b>
<b>Total</b>	\$150	\$90	\$54	\$81

So, even though you have won half the time, you have come out \$19 behind.

Surprisingly, it does not matter what order the wins and losses come in; if two are wins and two are losses, you come out behind. (You can check this.) So, because you are only going to win roughly half the time, and when you win half the time you actually lose money, it now seems to be a bad idea to gamble at Fogelin's Palace.

What should you do: gamble at Fogelin's Palace or not? Why?

## **EXPECTED OVERALL VALUE**

Given that lotteries usually have an extremely unfavorable expected monetary value, why do millions of people invest billions of dollars in them each year? Part of the answer is that some people are stupid, superstitious, or both. People will sometimes reason, "Somebody has to win; why not me?" They can also convince themselves that their lucky day has come. But that is not the whole story, for most people who put down money on lottery tickets realize that the bet is a bad bet, but think that it is worth doing anyway. People fantasize about what they will do with the money if they win, and fantasies are fun. Furthermore, if the bet is only \$1, and the person making the bet is not desperately poor, losing is not going to hurt much. Even if the expected monetary value on the lottery ticket is the loss of fifty cents, this might strike someone as a reasonable price for the fun of thinking about winning. (After all, you accept a sure loss of \$8 every time you pay \$8 to see a movie.) So a bet that is bad from a purely monetary point of view might be acceptable when other factors are considered.

The reverse situation can also arise: A bet may be unreasonable, even though it has a positive expected monetary value. Suppose, for example, that you are allowed to participate in a lottery in which a \$1 ticket gives you 1 chance in 10 million of getting a payoff of \$20 million. Here, as noted above, the expected monetary value of a \$1 bet is a profit of \$1, so from the point of view of expected monetary value, it is a good bet. This makes it sound reasonable to bet in this lottery, and a small bet probably is reasonable. But under these circumstances, would it be reasonable for you to sell everything you owned to buy lottery tickets? The answer to this is almost

certainly no, for, even though the expected monetary value is positive, the odds of winning are still low, and the loss of your total resources would be personally catastrophic.

When we examine the effects that success or failure will have on a particular person relative to his or her own needs, resources, preferences, and so on, we are then examining what we shall call the *expected overall value* or *expected utility* of a choice. Considerations of this kind often force us to make adjustments in weighing the significance of costs and payoffs. In the examples we just examined, the immediate catastrophic consequences of a loss outweigh the long-term gains one can expect from participating in the lottery.

Another factor that typically affects the expected overall value of a bet is the phenomenon known as the *diminishing marginal value* or *diminishing marginal utility* of a payoff as it gets larger. Suppose someone offers to pay a debt by buying you a hamburger. Provided that the debt matches the cost of a hamburger and you feel like having one, you might go along with this. But suppose this person offers to pay off a debt ten times larger by buying you ten hamburgers. The chances are that you will reject the offer, for even though ten hamburgers cost ten times as much as one hamburger, they are not worth ten times as much to you. At some point you will get stuffed and not want any more. After one or two hamburgers, the marginal value of one more hamburger becomes pretty low. The notion of marginal value applies to money as well. If you are starving, \$10 will mean a lot to you. You might be willing to work hard to get it. If you are wealthy, \$10 more or less makes little difference; losing \$10 might only be an annoyance.

Because of this phenomenon of diminishing marginal value, betting on lotteries is an even worse bet than most people suppose. A lottery with a payoff of \$20 million sounds attractive, but it does not seem to be twenty times more attractive than a payoff of \$1 million. So even if the expected monetary value of your \$1 bet in a lottery is the loss of \$0.50, the actual value to you is really something less than this, and so the bet is even worse than it seemed at first.

In general, then, when payoffs are large, the expected overall value of the payoff to someone is reduced because of the effects of diminishing marginal value. But not always. It is possible to think of exotic cases in which expected overall value increases with the size of the payoff. Suppose a witch told you that she would turn you into a toad if you did not give her \$10 million by tomorrow. You believe her, because you know for a fact that she has turned others into toads when they did not pay up. You have only \$1 to your name, but you are given the opportunity to participate in the first lottery described above, where a \$1 ticket gives you 1 chance in 20 million of hitting a \$10 million payoff. We saw that the expected monetary value of that wager was an unfavorable negative \$0.50. But now consider the overall value of \$1 to you if you are turned into a toad. Toads have no use for money, so to you, as a toad, the value of the dollar would drop to nothing. Thus, unless some other, more attractive alternatives are available, it would be reasonable to buy a lottery ticket, despite the unfavorable expected monetary value of the wager.

**EXERCISE III**

- Though the situation is somewhat far-fetched, suppose you are going to the drugstore to buy medicine for a friend who will die without it. You have only \$10—exactly what the medicine costs. Outside the drugstore a young man is playing three-card monte, a simple game in which the dealer shows you three cards, turns them over, shifts them briefly from hand to hand, and then lays them out, face down, on the top of a box. You are supposed to identify a particular card (usually the ace of spades); and, if you do, you are paid even money. You yourself are a magician and know the sleight-of-hand trick that fools most people, and you are sure that you can guess the card correctly nine times out of ten. First, what is the expected monetary value of a bet of \$10? In this context, would it be reasonable to make this bet? Why or why not?
- Provide an example of your own where a bet can be reasonable even though the expected monetary value is unfavorable. Then provide another example where the bet is unreasonable even though the expected monetary value is favorable. Explain what makes these bets reasonable or unreasonable.

**DISCUSSION QUESTION**

Consider the following game: You flip a coin continuously until you get tails once. If you get no heads (tails on the first flip), then you are paid nothing. If you get one heads (tails on the second flip), then you are paid \$2. If you get two heads (tails on the third flip), then you are paid \$4. If you get three heads, then you are paid \$8. And so on. The general rule is that for any number  $n$ , if you get  $n$  heads, then you are paid  $$2^n$ . What is the expected monetary value of this game? What would you pay to play this game? Why that amount rather than more or less?

**DECISIONS UNDER IGNORANCE**

So far we have discussed choices where the outcomes of the various options are not certain, but we know their probabilities. Decisions of this kind are called *decisions under risk*. In other cases, however, we do not know the probabilities of various outcomes. Decisions of this kind are called *decisions under ignorance* (or, sometimes, *decisions under uncertainty*). If we do not have any idea where the probabilities of various outcomes lie, the ignorance is complete. If we know that these probabilities lie within some general range, the ignorance is partial.

As an example of partial ignorance, suppose that, just after graduating from college, you are offered three jobs. First, the Exe Company offers you a salary of \$20,000. Exe is well established and secure. The next offer comes from the Wye Company. Here the salary is \$30,000, but Wye is a new company, so it is less secure. You think that this new company will probably do well, but you don't know how likely it is to last or for how long. Wye might go bankrupt, and then you will be left without a job. The final offer comes from the Zee Company, which is as stable as Exe and offers you a salary of \$40,000 per year. These offers are summarized in the following table:

	Wye does not go bankrupt	Wye goes bankrupt
Take job at Exe	\$20,000	\$20,000
Take job at Wye	\$30,000	\$0
Take job at Zee	\$40,000	\$40,000

Let's assume that other factors (such as benefits, vacations, location, interest, working conditions, bonuses, raises, and promotions) are all equally desirable in the three jobs. Which job should you take?

The answer is clear: Take the job from the Zee Company. This decision is easy because you end up better off regardless of whether or not Wye goes bankrupt, so it doesn't matter how likely Wye's bankruptcy is. Everyone agrees that you should choose any option that is best whatever happens. This is called the *rule of dominance*.

The problem with the rule of dominance is that it can't help you make choices when no option is better regardless of what happens. Suppose you discover that the letter from the Zee Company is a forgery—part of a cruel joke by your roommate. Now your only options are Exe and Wye. The job with Wye will be better if Wye does not go bankrupt, but the job with Exe will be better if Wye does go bankrupt. Neither job is better no matter what happens, so the rule of dominance no longer applies.

To help you choose between Exe and Wye, you might look for a rational way to assign probabilities despite your ignorance of which assignments are correct. One approach of this kind uses the *rule of insufficient reason*: When you have no reason to think that any outcome is more likely than any other, assume that the outcomes are equally probable. This assumption enables us to calculate expected monetary value or utility, as in the preceding sections, and then we can choose the option with the highest expected utility. In our example, this rule of insufficient reason favors the job at Exe, because your expected income in that job is \$20,000, whereas your expected income in the job at Wye is only \$15,000 ( $= 0.5 \times \$30,000$ ), assuming that the Wye company has as much chance of going bankrupt as of staying in business.

The problem with the rule of insufficient reason is that it may seem arbitrary to assume that unknown probabilities are equal. Often we suspect that the probabilities of various outcomes are not equal, even while we do not know what the probabilities are. Moreover, the rule of insufficient reason yields different results when the options are described differently. We can distinguish four possibilities: Wye goes bankrupt, Wye stays the same size, Wye increases in size, and Wye decreases in size but stays in business. If we do not have any reason to see any of these outcomes as more likely than any other, then the rule of insufficient reason tells us to assign them equal probabilities. On that assumption, and if you will keep your job as long as Wye stays in business, then you have only one chance in four of losing your job; so your expected income in the job at Wye is now \$22,500 ( $= 0.75 \times \$30,000$ ). Thus, if we stick with the rule of insufficient reason, the expected value of the job at Wye and whether you should take that job seem to depend on how the options are divided. That seems crazy in this case.

Another approach tries to work without any assumptions about probability in cases of ignorance. Within this approach, several rules might be adopted. One possibility is the *maximax rule*, which tells you to choose the option whose best outcome is better than the best outcome of any other option. If you follow the maximax rule, then you will accept the job with the Wye Company, because the best outcome of that job is a salary of \$30,000 when this new company does not go bankrupt, and this is better than any outcome with the Exe Company. Optimists and risk takers will favor this rule.

Other people are more pessimistic and tend to avoid risks. They will favor a rule more like the *maximin rule*, which says to choose the option whose worst outcome is better than the worst outcome of any other option. If you follow the maximin rule, you will accept the job with the Exe Company, because the worst outcome in that job is a steady salary of \$20,000, whereas the worst outcome is unemployment if you accept the job with the Wye Company.

Each of these rules works by focusing exclusively on part of your information and disregarding other things that you know. The maximax rule considers only the best outcomes for each option—the best-case scenario. The maximin rule pays attention to only the worst outcome for each option—the worst-case scenario. Because they ignore other outcomes, the maximax rule strikes many people as too risky (since it does not consider how much you could lose by taking a chance), and the maximin rule strikes many people as too conservative (since it does not consider how much you could have gained if you had taken a small risk).

Another problem is that the maximax and maximin rules do not take probabilities into account at all. This makes sense when you know nothing about the probabilities. But when some (even if limited) information about probabilities is available, then it seems better to use as much information as you have. Suppose, for example, that each of two options might lead to disaster, and you do not know how likely a disaster is after either option, but you do know that one option is more likely to lead to disaster than another. In such situations, some decision theorists argue that you should choose the option that minimizes the chance that any disaster will occur. This is called the *disaster avoidance rule*.

To illustrate this rule, consider a different kind of case:

A forty-year-old man is diagnosed as having a rare disease and consults the world's leading expert on the disease. He is informed that the disease is almost certainly not fatal but often causes serious paralysis that leaves its victims bedridden for life. (In other cases it has no lasting effects.) The disease is so rare that the expert can offer only a vague estimate of the probability of paralysis: 20 to 60 percent. There is an experimental drug that, if administered now, would almost certainly cure the disease. However, it kills a significant but not accurately known percentage of those who take it. The expert guesses that the probability of the drug being fatal is less than 20 percent, and the patient thus assumes that he is definitely less likely to die if he takes the drug than he is to be paralyzed if he lets the disease run its course. The patient would regard bedridden life as preferable to death, but he considers both outcomes as totally disastrous compared to continuing his life in good health. Should he take the drug?<sup>2</sup>

Since the worst outcome is death, and this outcome will not occur unless he takes the drug, the maximin rule would tell him not to take the drug. In contrast, the disaster avoidance rule would tell him to take the drug, because both death and paralysis are disasters and taking the drug minimizes his chances that any disaster will occur. Thus, although the disaster avoidance rule opposes risk taking, it does so in a different way than the maximin rule.

We are left, then, with a plethora of rules: dominance, insufficient reason, maximax, maximin, and disaster avoidance. Other rules have been proposed as well. With all of these rules in the offing, it is natural to ask which is correct. Unfortunately, there is no consensus on this issue. Each rule applies and seems plausible in some cases but not in others. Many people conclude that each rule is appropriate to different kinds of situations. It is still not clear, however, which rule should govern decisions in which circumstances. The important problem of decision under ignorance remains unsolved.

**DISCUSSION QUESTIONS**

1. In the game of ignorance, you draw one card from a deck, but you do not know how many cards or which kinds of cards are in the deck. It might be a normal deck or it might contain only diamonds or only aces of spades or any other combination of cards. It costs nothing to play. If you bet that the card you draw will be a spade, and it is a spade, then you win \$100. If you bet that the card you draw will not be a spade, and it is not a spade, then you win \$90. You may make only one bet. Which bet would you make if you followed the maximax rule? The maximin rule? The disaster avoidance rule? The rule of insufficient reason? Which rule seems most plausible this case? Which bet should you make? Why?
2. In which circumstances do you think it is appropriate to use the dominance rule? The rule of insufficient reason? The maximax rule? The maximin rule? The disaster avoidance rule? Why?
3. Suppose that you may choose either of two envelopes. You know that one envelope contains twice as much money as the other, but you do not know the amount of money in either envelope. You choose an envelope, open it, and see that it contains \$100. Now you know that the other envelope must contain either \$50 or \$200. At this point, you are given a choice: You may exchange your envelope for the other envelope. Should you switch envelopes, according to the rule of insufficient reason? Is this result plausible? Why or why not?

**NOTES**

<sup>1</sup> If the lottery gave a consolation prize of a shiny new quarter to all losers, their net loss would be only seventy-five cents. Since most lotteries do not give consolation prizes, the net loss equals the cost of playing such lotteries.

<sup>2</sup> Gregory Kavka, "Deterrence, Utility, and Rational Choice," reprinted in *Moral Paradoxes of Nuclear Deterrence* (New York: Cambridge University Press, 1987), 65–66. Kavka uses this medical example to argue for his disaster avoidance rule and, by analogy, to defend the rationality of nuclear deterrence.

## FALLACIES

*When inferences are defective, they are called fallacious. When defective styles of reasoning are repeated over and over, because people often get fooled by them, then we have an argumentative fallacy that is worth flagging with a name. The number and variety of argumentative fallacies are limited only by the imagination. Consequently, there is little point in trying to construct a complete list of fallacies. What is crucial is to get a feel for the most common and most seductive kinds of fallacy. Once this is done, we should be able to recognize many other kinds as well. The goal of Part IV is to develop that skill.*



## FALLACIES OF VAGUENESS

This chapter examines one of the main ways in which arguments can be defective or fallacious because language is not used clearly enough for the context. This kind of unclarity is vagueness. Vagueness occurs when, in a given context, a term is used in a way that allows too many cases in which it is unclear whether or not the term applies. Vagueness underlies several common fallacies, including three kinds of slippery-slope arguments.

### USES OF UNCLARITY

In a good argument, a person states a conclusion clearly and then, with equal clarity, gives reasons for this conclusion. The arguments of everyday life often fall short of this standard. Usually, unclear language is a sign of unclear thought. There are times, however, when people are intentionally unclear. They might use unclarity for poetic effect or to leave details to be decided later. But often their goal is to confuse others. This is called *obfuscation*.

Before we look at the various ways in which language can be unclear, a word of caution is needed: There is no such thing as absolute clarity. Whether something is clear or not depends on the context in which it occurs. A botanist does not use common vocabulary in describing and classifying plants. At the same time, it would usually be foolish for a person to use botanical terms in describing the appearance of his or her backyard. Aristotle said that it is the mark of an educated person not to expect more rigor than the subject matter will allow. Because clarity and rigor depend on context, it takes judgment and good sense to pitch an argument at the right level.

#### Non Sequitur

by Wiley



NON SEQUITUR © 1995 Wiley Ink, Inc.  
Dist. by UNIVERSAL UCLICK. Reprinted with  
permission. All rights reserved.

## VAGUENESS

Perhaps the most common form of unclarity is *vagueness*. It arises when a concept applies along a continuum or a series of very small changes. The standard example is baldness. A person with a full head of hair is not bald. A person without a hair on his head is bald. In between, however, is a range of cases in which we cannot say definitely whether the person is bald or not. These are called *borderline cases*. Here we say something less definite, such as that this person is “going bald.”

Our inability to apply the concept of baldness in a borderline case is not due to ignorance of the number of hairs on the person’s head. It will not help to count the number of hairs there. Even if we knew the exact number, we would still not be able to say whether the person was bald or not. The same is true of most adjectives that concern properties admitting of degrees—for example, “rich,” “healthy,” “tall,” “wise,” and “ruthless.”

For the most part, imprecision—the lack of sharply defined limits—causes little difficulty. In fact, this is a useful feature of our language, for suppose we *did* have to count the number of grains of salt between our fingers to determine whether or not we hold a *pinch* of salt. It would take a long time to follow a simple recipe that calls for a pinch of salt.

Yet, difficulties can arise when borderline cases themselves are at issue. Suppose that a state passes a law forbidding all actions in public that are obscene. There will be many cases that clearly fall under this law and many cases that clearly do not fall under it. There will also be many cases in which it will not be clear whether or not they fall under this law. Laws are sometimes declared unconstitutional for this very reason. Here we shall say that the law is vague. In calling the law vague, we are criticizing it. We are not simply noticing the existence of borderline cases, for there will usually be borderline cases no matter how careful we are. Instead, we are saying that there are *too many* borderline cases for this context. More precisely, we shall say that an expression in a given context is *used vaguely* if it leaves open too wide a range of borderline cases for the successful and legitimate use of that expression in that context.

Vagueness thus depends on context. To further illustrate this context dependence, consider the expression “light football player.” There are, of course, borderline cases between those football players who are light and those who are not light. But on these grounds alone, we would not say that the expression is vague. It is usually a perfectly serviceable expression, and we can indicate borderline cases by saying such things as “Jones is a bit light for a football player.” Suppose, however, that Ohio State and Cal Tech wish to have a game between their light football players. It is obvious that the previous understanding of what counts as being light is too vague for this new context. At Ohio State, anyone under 210 pounds is considered light. At Cal Tech, anyone over 150 pounds is considered heavy. What is needed, then, is a ruling, such as that anyone under 175 pounds will be considered

a lightweight. This example illustrates a common problem and its solution. A term that works perfectly well in one area becomes vague when applied in some other (usually more specialized) area. This vagueness can then be removed by adopting more precise rules in the problematic area. Vagueness is resolved by definition.

### EXERCISE I

For each of the following terms, give one case to which the term clearly applies, one case to which the term clearly does not apply, and one borderline case. Then try to explain why the borderline case is a borderline case.

Example: In the northern hemisphere, “summer month” clearly applies to July; clearly does not apply to January; and June is a borderline case, because the summer solstice is June 21, and schools usually continue into June, but June, July, and August are, nonetheless, often described as the summer months.

1. large animal
2. populous country
3. long book
4. old professor
5. popular singer
6. powerful person
7. difficult subject
8. late meeting
9. arriving late to a meeting

### EXERCISE II

Each of the following sentences contains words or expressions that are potentially vague. Describe a context in which this vagueness might make a difference, and explain what difference it makes. Then reduce this vagueness by replacing the italicized expression with one that is more precise.

Example: Harold has a *bad* reputation.

Context: If Harold applies for a job as a bank security guard, then some but not all kinds of bad reputation are relevant. A reputation for doing bad construction work is irrelevant, but a reputation for dishonesty is relevant.

Replacement: Harold is a known thief.

1. Ross has a *large* income.
2. Cocaine is a *dangerous* drug.
3. Ruth is a *clever* woman.
4. Andre is a *terrific* tennis player.

(continued)

5. Mark is not doing *too well* (after his operation).
6. Shaq's a *big* fellow.
7. Dan's grades are *low*.
8. Walter can't see *well*.
9. The earthquake was a *disaster*.
10. The news was *wonderful*.

## HEAPS

---

The existence of borderline cases is essential to various styles of reasoning that have been identified and used since ancient times. One such argument was called the *argument from the heap* or the *sorites argument* (from the Greek word "soros," which means "heap"). The classic example was intended to show that it is impossible to produce a heap of sand by adding one grain at a time. As a variation on this, we will show that no one can become rich. The argument can be formulated as a long series like this:

- (1) Someone with only one cent is not rich.
  - (2) If someone with only one cent is not rich, then someone with only two cents is not rich.
  - ∴ (3) Someone with only two cents is not rich. (from 1-2)
  - (4) If someone with only two cents is not rich, then someone with only three cents is not rich.
  - ∴ (5) Someone with only three cents is not rich. (from 3-4)
  - (6) If someone with only three cents is not rich, then someone with only four cents is not rich.
  - ∴ (7) Someone with only four cents is not rich. (from 5-6)
- [and so on, until:]
- ∴ (199,999,999,999) Someone with only 100,000,000,000 cents is not rich.

The problem, of course, is that someone with 100,000,000,000 cents *is* rich. If someone denies this, we can keep on going. Or we can just sum up the whole argument like this:

- (1\*) Someone with only one cent is not rich.
- (2\*) For any number,  $n$ , if someone with only  $n$  cents is not rich, then someone with  $n + 1$  cents is not rich.
- ∴ (3\*) Someone with any number of cents is not rich.

Premise (2\*) is, of course, just a generalization of premises (2), (4), (6), and so on.

Despite its plausibility, everyone should agree that there is something wrong with this argument. If we hand over enough pennies to Peter, previously poor Peter will become the richest person in the world. Another sign of a problem is that a parallel argument runs in the other direction: Someone with 100 billion cents is rich. For any number,  $n$ , if someone with  $n$  cents is rich, then someone with  $n - 1$  cents is also rich. Therefore, someone with no cents at all is rich. This is absurd (since we are not talking about how rich one's life can be as long as one has friends).

We can see that these arguments turn on borderline cases in the following way: The argument would fail if we removed borderline cases by laying down a ruling (maybe for tax purposes) that anyone with a million dollars or more is rich and anyone with less than this is not rich. A person with \$999,999.99 would then pass from not being rich to being rich when given a single penny, so premise (2\*) would be false at that point under this ruling. Of course, we do not usually use the word "rich" with this much precision. We see some people as clearly rich and others as clearly not rich, but in between there is a fuzzy area where we are not prepared to say that people either are or are not rich. In this fuzzy area, as well as in the clear areas, a penny one way or the other will make no difference.

That is how the argument works, but exactly where does it go wrong? This question is not easy to answer and remains a subject of vigorous debate. Here is one way to view the problem: Consider a person who is 80 pounds overweight, where we would all agree that that person would pass from being fat to not being fat by losing over 100 pounds. If he or she lost an ounce a day for five years, this would be *equivalent* to losing just over 114 pounds. An argument from the heap denies that this person would ever cease to be fat. (So what is the point of dieting?) Anyone who accepted that conclusion, or (3\*), would seem to claim that a series of insignificant changes *cannot* be equivalent to a significant change. Surely, this is wrong. Here we might be met with the reply that every change must occur at some particular time (and place), but there would be no particular day on which this person would pass from being fat to not being fat. The problem with this reply is that, with concepts like this, changes seem to occur gradually over long stretches of time without occurring at any single moment. Anyway, however or whenever it occurs, the change does occur. Some people do cease to be fat if they lose enough weight.

This tells us that conclusions of arguments from the heap, such as (3\*), are false, so these arguments cannot be sound. Almost everyone agrees to that much. Moreover, if an appropriate starting point is chosen, then premises like (1) and (1\*) will also be accepted as true by almost everyone. So the main debate focuses on premise (2\*) and on whether the argument is valid. Some philosophers reject premise (2\*) and claim that there is a precise point at which a person becomes rich, even though we don't know where that point is. Others try to avoid any sharp cutoff point by developing some kind of alternative logic. Still others just admit that the premises seem true, and the argument seems valid, but the conclusion seems false, so the argument

creates a paradox to which they have no solution. These views become complicated and technical, so we will not discuss them here. Suffice it to say that almost everyone agrees that conclusions like (3\*) and (199,999,999,999) are false, so arguments from the heap are unsound for one reason or another. That is why such arguments are labeled *fallacies*.

#### DISCUSSION QUESTION

Where exactly do *you* think arguments from the heap go astray?

### SLIPPERY SLOPES

Near cousins to arguments from the heap are *slippery-slope arguments*, but they reach different conclusions. Whereas heap arguments conclude that nothing has a certain property, such as baldness, a slippery-slope argument could be trotted out to try to show that there is no *real* or *defensible* or *significant* or *important* difference between being bald and not being bald. The claim is not that no change occurs because the person who loses all his hair is still not bald, as in an argument from the heap. Instead, the slippery-slope argument claims that we should not classify people as either bald or not bald, because there is no significant difference between these classifications.

Whether a difference is significant depends on a variety of factors. In particular, what is significant for one purpose might not be significant for other purposes. Different concerns then yield different kinds of slippery-slope arguments. We will discuss three kinds, beginning with conceptual slippery-slope arguments.

#### CONCEPTUAL SLIPPERY-SLOPE ARGUMENTS

Conceptual slippery-slope arguments try to show that things at opposite ends of a continuum do not differ in any way that would be important enough to justify drawing a distinction in one's concepts or theories. Such arguments often seem to depend on the following principles:

1. We should not draw a distinction between things that are not significantly different.
2. If *A* is not significantly different from *B*, and *B* is not significantly different from *C*, then *A* is not significantly different from *C*.

This first principle is interesting, complicated, and at least *generally* true. We shall examine it more closely in a moment. The second principle is obviously false. As already noted, a series of insignificant differences can add up to a

significant difference. U.S. Senator Everett Dirksen put the point memorably when he said, “A billion dollars here and a billion dollars there can add up to some real money.” To the extent that conceptual slippery-slope arguments depend on this questionable assumption, they provide no more support for their conclusions than do arguments from the heap.

Unlike arguments from the heap, however, conceptual slippery-slope arguments do often lead people to accept their conclusions. Slippery-slope arguments have been used to deny the difference between sanity and insanity (some people are just a little weirder than others) and between amateur and professional athletics (some athletes just get paid a little more or more directly than other athletes). When many small differences make a big difference, such conceptual slippery-slope arguments are fallacious.

This fallacy is seductive, because it is often hard to tell when many small differences *do* make a big difference. Here is a recent controversial example: Some humans have very dark skin. Others have very pale skin. As members of these different groups marry, their children’s skin can have any intermediate shade of color. This smooth spectrum leads some people to deny that any differences among races will be important to developed theories in biology. Their argument seems to be that the wealth of intermediate cases will make it difficult or impossible to formulate precise and exceptionless laws that apply to one racial group but not to others, so differences among races will play no important role in sciences that seek such laws. Critics respond that some scientific laws about races still might hold without exception even if skin color and other features do vary in tiny increments.

Whichever side one takes, this controversy shows that, even if there is a smooth spectrum between end points, this continuity is not enough *by itself* to show that there are no scientifically significant differences among races. That conclusion would need to be supported by more than just a conceptual slippery-slope argument. To show that certain concepts are useless for the purposes of a certain theory, one would need to add more information, particularly about the purposes of that theory and its laws. That is what determines which differences are important in that particular area. Conceptual slippery-slope arguments might work in conjunction with such additional premises, but they cannot work alone.

### EXERCISE III

Whenever we find one thing passing over into its opposite through a gradual series of borderline cases, we can construct (a) an argument from the heap and (b) a conceptual slippery-slope argument by using the following method: Find some increase that will not be large enough to carry us outside the borderline area. Then use the patterns of argument given above. Applying this method,

(continued)

formulate arguments for the following claims. Then explain what is wrong with these arguments.

1. a. There are no heaps.  
b. There is no difference between a heap and a single grain of sand.
2. a. Nobody is tall.  
b. There is no difference between being tall and being short.
3. a. Books do not exist.  
b. There is no difference between a book and a pamphlet.
4. a. Heat is not real.  
b. There is no difference between being hot and being cold.
5. a. Taxes are never high.  
b. There is no difference between high taxes and low taxes.
6. a. Science is an illusion.  
b. There is no difference between science and faith.

#### DISCUSSION QUESTIONS

1. Do you think that differences among races have any role in developed theories in biology or sociology or any other science? Why or why not?
2. If animals evolve gradually from one species to another, does that show that there is no significant difference in biology between any species (say, horses and dogs)? Why or why not? Does it show that there is no important difference in moral theory between the rights of humans and the rights of animals in other species? Why or why not?

### FAIRNESS SLIPPERY-SLOPE ARGUMENTS

When borderline cases form a continuum, if someone classifies a case at one end of the continuum, an opponent often challenges this classification by asking, "Where do you draw the line?" This rhetorical question suggests that differences along the continuum are "*just* a matter of degree," so it is arbitrary and, hence, unfair to draw a line at any particular point along the continuum, because then very similar cases on different sides of the line will be classified and treated very differently.

Questions about the fairness of drawing a line often arise in the law. For example, given reasonable cause, the police generally do not have to obtain a warrant to search a motor vehicle, for the obvious reason that the vehicle might be driven away while the police go to a judge to obtain a warrant. On the other hand, with few exceptions, the police may not search a

person's home without a search warrant. In the case of *California v. Carney*,<sup>1</sup> the U.S. Supreme Court had to rule on whether the police needed a warrant to search for marijuana in an "oversized van, fully mobile," parked in a downtown parking lot in San Diego. Because the van was a fully mobile vehicle, it seemed to fall under the first principle; but because it also served as its owner's home, it seemed to fall under the second. The difficulty, as the Court saw, was that there is a gray area between those things that clearly are motor vehicles and not homes (for example, motorcycles) and those things that clearly are homes and not motor vehicles (for example, apartments). Chief Justice Warren Burger wondered about a mobile home in a trailer park hooked up to utility lines with its wheels removed. Justice Sandra Day O'Connor asked whether a tent, because it is highly mobile, could also be searched without a warrant. As the discussion continued, houseboats (with or without motors or oars), covered wagons, and finally a house being moved from one place to another on a trailer truck came under examination. In the end, our highest court decided that the van in question was a vehicle and could be searched without first obtaining a warrant to do so. The court did not fully explain why it is fair to allow warrantless searches—and to send people to jail as a result—in cases of vans used as homes but not in other very similar cases.

Questions about where to draw a line often have even more important implications than warrantless searches. Consider the death penalty. Most societies have reserved the death penalty for those crimes they consider the most serious. But where should we draw the line between crimes punishable by death and crimes not punishable by death? Should the death penalty be given to murderers of prison guards? To rapists? To drug dealers? To drunk drivers who cause death? Wherever we draw the line, it seems to be an unavoidable consequence of the death penalty that similar cases will be treated in radically different ways. A defender of the death penalty can argue that it is not unfair to draw a line because, once the line is drawn, the public will have fair warning about which crimes are subject to the death penalty and which are not. It will then be up to each person to decide whether to risk his or her life by crossing this line. It remains a matter of debate, however, whether the law can be administered in a predictable way that makes this argument plausible.

The finality of death raises a profoundly difficult problem in another area too: the legalization of abortion. There are some people who think abortion is never justified and ought to be declared totally illegal. There are others who think abortion does not need any justification at all and should be completely legalized. Between these extremes, there are many people who believe abortion is justified in certain circumstances but not in others (such as when abortion is the only way to save the life of the mother but not when it prevents only lesser harms to the mother). There are also those who think abortion should be allowed for a certain number of months of pregnancy, but not thereafter. People holding these middle positions face the problem of deciding where to draw a line, and this makes them subject to criticisms from holders of either extreme position.

This problem admits of no easy solution. Because every line we draw will seem arbitrary to some extent, a person who holds a middle position needs to argue that it is better to draw *some* line—even a somewhat arbitrary one—than to draw no line at all. The recognition that some line is needed, and why, can often help us locate the real issues. This is the first step toward a reasonable position.

Of course, this still does not tell us *where* to draw the line. A separate argument is needed to show that the line should be drawn at one point, or in one area, rather than another. In the law, such arguments often appeal to value judgments about the effects of drawing the line at one place rather than another. For example, it is more efficient to draw a line where it is easy to detect, and drawing the line at one place will provide greater protection for some values or some people than drawing it at another place. Different values often favor drawing different lines, and sometimes such arguments are not available at all. Thus, in the end, it will be difficult to solve many of these profound and important problems.

#### DISCUSSION QUESTION

Is it unfair for teachers to fail students who get one point out of a hundred less than others students who pass? Why or why not? Would an alternative grading system be fairer?

## Farcus

by David Waisglass  
Gordon Coulthart



1992 Farcus Cartoons/LaughingStock Licensing Inc.

**What began with a few pencils  
and paper clips ...**

## CAUSAL SLIPPERY-SLOPE ARGUMENTS

Another common kind of argument is also often described as a slippery-slope argument. In these arguments, the claim is made that, once a certain kind of event occurs, other similar events will also occur, and this will lead eventually to disaster. The most famous (or infamous) argument of this kind was used by the U.S. government to justify its intervention in Vietnam in the 1960s. It was claimed that, if the communists took over Vietnam, they would then take over Cambodia, the rest of Asia, and other continents, until they ruled the whole world. This was called the *domino theory*, since the fall of one country would make neighboring countries fall as well. Arguments of this kind are sometimes called *domino arguments*. Such arguments claim that one event, which might not seem bad by itself, would lead to other, more horrible events, so such arguments can also be called *parades of horrors*.

Causal slippery slopes can also slide into good results. After all, someone who wants communists to take over the world might use the above domino argument to show why the United States should *not* intervene in Vietnam. Such optimistic slippery-slope arguments are, however, much less common than parades of horrors, so we will limit our discussion to the pessimistic versions.

These arguments resemble other slippery-slope arguments in that they depend on a series of small changes. The domino argument does not, however, claim that there is no difference between the first step and later steps—between Vietnam going communist and the rest of Asia going communist. Nor is there supposed to be anything unfair about letting Vietnam go communist without letting other countries also go communist. The point of a parade of horrors is that certain events will *cause* horrible effects because of their similarity or proximity to other events. Since the crucial claim is about causes and effects, these arguments will be called *causal slippery-slope arguments*.

We saw another example in Chapter 4. While arguing against an increase in the clerk hire allowance, Kyl says,

The amount of increase does not appear large. I trust, however, there is no one among us who would suggest that the addition of a clerk would not entail allowances for another desk, another typewriter, more materials, and it is not beyond the realm of possibility that the next step would then be a request for additional office space, and ultimately new buildings.<sup>2</sup>

Although this argument is heavily guarded, the basic claim is that increasing the clerk hire allowance is likely to lead to much larger expenditures that will break the budget. The argument can be represented more formally this way:

- (1) If the clerk hire allowance is increased, other expenditures will also probably be increased.
  - (2) These other increases would be horrible.
- ∴ (3) The clerk hire allowance should not be increased.

Opponents can respond in several ways. One response is to deny that the supposedly horrible effects really are so horrible. One might argue, for example, that additional office space and new buildings would be useful. This response is often foreclosed by describing the effects in especially horrible terms.

A second possible response would be to deny that increasing the clerk hire allowance really would have the horrible effects that are claimed in the first premise. One might argue, for example, that the old offices already have adequate room for additional clerks.

Often the best response is a combination of these. One can admit that certain claimed effects would be horrible, but deny that these horrible effects really are likely. Then one can acknowledge that some more minor problems will ensue, but argue that these costs are outweighed by the benefits of the program.

To determine which, if any, of these responses is adequate, one must look closely at each particular argument and ask the following questions:

Are any of the claimed effects really very bad?

Are any of these effects really very likely?

Do these dangers outweigh all the benefits of what is being criticized?

If the answers to all these questions are “Yes,” then the causal slippery-slope argument is strong. But if any of these questions receives a negative answer, then the causal slippery-slope argument is questionable on that basis.

#### EXERCISE IV

Classify each of the following arguments as either (H) an argument from the heap, (C) a conceptual slippery-slope argument, (F) a fairness slippery-slope argument, or (S) a causal slippery-slope argument. Explain why you classify each example as you do.

1. We have to take a stand against sex education in junior high schools. If we allow sex education in the eighth grade, then the seventh graders will want it, and then the sixth graders, and pretty soon we will be teaching sex education to our little kindergartners.
2. People are found not guilty by reason of insanity when they cannot avoid breaking the law. But people who are brought up in certain deprived social circumstances are not much more able than the insane to avoid breaking the law. So it would be unjust to find them guilty.
3. People are called mentally ill when they do very strange things, but many so-called eccentrics do things that are just as strange. So there is no real difference between insanity and eccentricity.
4. If you try to smoke one cigarette a day, you will end up smoking two and then three and four and five, and so on, until you smoke two packs every day. So don't try even one.

(continued)

5. A human egg one minute after fertilization is not very different from what it is one minute later, or one minute after that, and so on. Thus, there is really no difference between just-fertilized eggs and adult humans.
6. Since no moment in the continuum of development between an egg and a baby is especially significant, it is not fair to grant a right to life to a baby unless one grants the same right to every fertilized egg.
7. If we let doctors kill dying patients who are in great pain, then they will kill other patients who are in less pain and patients who are only slightly disabled. Eventually, they will kill anyone who is not wanted by society.

**EXERCISE V**

Explain the reasons, if any, for drawing a definite line in each of the following cases. Then further explain how this line can be drawn, if at all, in a reasonable way.

1. Minimum (or maximum?) age to drive a car
2. Minimum age to vote
3. Minimum age to enter (or be drafted into) the military
4. Minimum age to drink alcoholic beverages
5. Minimum age for election to the presidency
6. Maximum age before retirement becomes mandatory

**EXERCISE VI**

Determine whether each of the following arguments provides adequate support, or any support, for its conclusion. Explain why.

1. I shouldn't get a speeding ticket for going fifty-six miles per hour, because my driving did not all of a sudden get more dangerous when I passed the speed limit of fifty-five.
2. No student should ever be allowed to ask a question during a lecture, because once one student asks a question, then another one wants to ask a question, and pretty soon the teacher doesn't have any time left to lecture.
3. Pornography shouldn't be illegal, because you can't draw a line between pornography and erotic art.
4. Marijuana should be legal, because it is no more dangerous than alcohol or nicotine.

*(continued)*

5. Marijuana should be illegal, because people who try marijuana are likely to go on to try hashish, and then snorting cocaine, and then freebasing cocaine or shooting heroin.
6. The government should not put any new restrictions on free trade, because once they place some restrictions, they will place more and more until foreign trade is so limited that our own economy will suffer.
7. Governments should never bargain with any terrorist. Once they do, they will have to bargain with every other terrorist who comes along.
8. If assault weapons are banned, Congress will ban handguns next, and then rifles. Eventually, hunters will not be able to hunt, and law-abiding citizens will have no way to defend themselves against criminals.

#### DISCUSSION QUESTIONS

1. Explain and evaluate the following argument against restrictions on hate speech:

To attempt to craft free speech exceptions only for racist speech would create a significant risk of a slide down the proverbial “slippery slope.” . . . Censorial consequences could result from many proposed or adopted university policies, including the Stanford code, which sanctions speech intended to “insult or stigmatize” on the basis of race or other prohibited grounds. For example, certain feminists suggest that all heterosexual sex is rape because heterosexual men are aggressors who operate in a cultural climate of pervasive sexism and violence against women. Aren’t these feminists insulting or stigmatizing heterosexual men on the basis of their sex and sexual orientation? And how about a Holocaust survivor who blames all (“Aryan”) Germans for their collaboration during World War II? Doesn’t this insinuation insult and stigmatize on the basis of national and ethnic origin? And surely we can think of numerous other examples that would have to give us pause.<sup>3</sup>

2. Explain and evaluate the following response to critics of college restrictions on hate speech:

[Defenders of such restrictions] will ask whether an educational institution does not have the power . . . to enact reasonable regulations aimed at assuring equal personhood on campus. If one characterizes the issue this way, . . . a different set of slopes will look slippery. If we do not intervene to protect equality here, what will the next outrage be?<sup>4</sup>

3. When John Stewart interviewed William Bennett (former Secretary of Education under President Ronald Reagan) about gay marriage, both of them used slippery slopes and responded to each other’s slippery slopes in the following exchange. What kinds of slippery slopes did they use? Was either argument better than the other? Was either response better than the other? Why or why not?

**BENNETT:** The question is: How do you define marriage? Where do you draw the line? What do you say to the polygamist?

**STEWART:** You don't say anything to the polygamist. That is a choice, to get three or four wives. That is not a biological condition that "I gots to get laid by different women that I am married to." That's a choice. Being gay is part of the human condition. There's a huge difference.

**BENNETT:** Well, some people regard their human condition as having three women. Look, the polygamists are all over this.

**STEWART:** Then let's go slippery slope the other way. If government says I can define marriage as between a man and a woman, what says they can't define it between people of different income levels, or they can decide whether or not you are a suitable husband for a particular woman?

**BENNETT:** Because gender *matters* in marriage, it has mattered to every human society, it matters in every religion . . .

**STEWART:** Race matters in every society as well. Isn't progress understanding?<sup>5</sup>

4. What, if anything, is shown when slippery-slope arguments can be used on both sides of an issue?

## NOTES

<sup>1</sup> 1471 U.S. 386 (1984). This case was reported by Linda Greenhouse, "Of Tents with Wheels and Houses with Oars," *New York Times*, May 15, 1985.

<sup>2</sup> *Congressional Record* (vol. 107, part 3, March 15, 1961, pp. 4059–60).

<sup>3</sup> Nadine Strossen, "Regulating Racist Speech on Campus: A Modest Proposal?" *Duke Law Journal* 39 (1990): 537–38. When she wrote this, Strossen was on the National Board of Directors of the American Civil Liberties Union.

<sup>4</sup> Richard Delgado, "Campus Antiracism Rules: Constitutional Narratives in Collision," *Northwestern University Law Review* 85 (1991): 346.

<sup>5</sup> *The Daily Show* with John Stewart, Episode 11069 on 06/06/06 on Comedy Central.



## FALLACIES OF AMBIGUITY

This chapter examines fallacies that arise from a second kind of unclarity: ambiguity. Ambiguity occurs when it is unclear which meaning of a term is intended in a given context. Ambiguity leads to the fallacy of equivocation, which will be defined and illustrated. The chapter closes with a discussion of different kinds of definitions that can be useful in avoiding or responding to fallacies of clarity.

### AMBIGUITY

---

The idea of vagueness is based on a common feature of words in our language: Many of them leave open a range of borderline cases. The notion of *ambiguity* is also based on a common feature of our language: Words often have a number of different meanings. For example, the *New Merriam-Webster Pocket Dictionary* has the following entry under the word “cardinal”:

- cardinal** adj. 1: of basic importance; chief, main, primary,
  - 2: of cardinal red color.
- n. 1: an ecclesiastical official of the Roman Catholic Church ranking next below the pope,
  - 2: a bright red,
- 3: any of several American finches of which the male is bright red.

In the plural, “the Cardinals” is the name of an athletic team that inhabits St. Louis; “cardinal” also describes the numbers used in simple counting.

It is not likely that people would get confused about these very different meanings of the word “cardinal,” but we might imagine a priest, a bird-watcher, and a baseball fan all hearing the remark, “The cardinals are in town.” The priest would prepare for a solemn occasion, the bird-watcher would get out binoculars, and the baseball fan would head for the stadium. In this context, the remark might be criticized as *ambiguous*. More precisely, we shall say that an expression in a given context is used *ambiguously* if and only if it is misleading or potentially misleading because it is hard to tell which of a number of possible meanings is intended in that context.

Using this definition, the word “bank” is *not* used ambiguously in the following sentence:

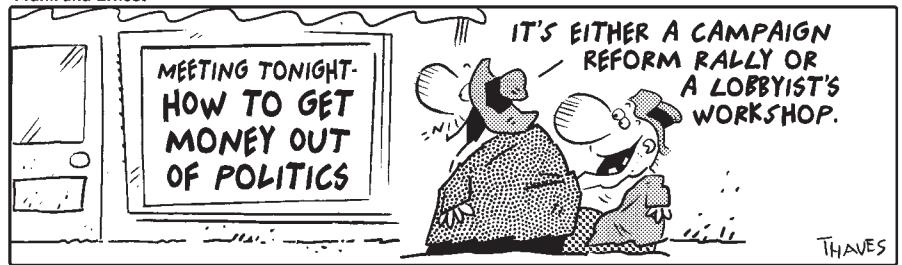
Joan deposited \$500 in the bank and got a receipt.

Some writers, however, call an expression ambiguous simply if it admits of more than one interpretation, without adding that it is not possible to tell which meaning is intended. With this definition, the above sentence is ambiguous because it could mean that Joan placed \$500 in a riverbank, and someone, for whatever reason, gave her a receipt for doing so. On this second definition of ambiguity, virtually every expression is ambiguous, because virtually every expression admits of more than one interpretation. On our first definition, only uses of expressions that are misleading or potentially misleading will be called ambiguous. In what follows, we will use the word “ambiguous” in accordance with the first definition. Ambiguity then depends on the context, because whether something is misleading also depends on context.

In everyday life, context usually settles which of a variety of meanings is appropriate. Yet sometimes genuine misunderstandings do arise. An American and a European discussing “football” may have different games in mind. The European is talking about what Americans call “soccer”; the American is talking about what Europeans call “American football.” It is characteristic of the ambiguous use of a term that when it comes to light, we are likely to say something like, “Oh, you mean *that* kind of cardinal!” or “Oh, you were talking about *American* football!” This kind of misunderstanding can cause trouble. When it does, if we want to criticize the expression that creates the problem, we call it ambiguous.

Thus, “ambiguous” is both dependent on context and a term of criticism in much the same ways as “vague.” But these kinds of unclarity differ in other ways. In a context where the use of a word is ambiguous, it is not clear which of two meanings to attach to a word. In a context where the use of a word is vague, we cannot attach *any* precise meaning to the use of a word.

**Frank and Ernest**



© 2000 Thaves. Reprinted with permission. Newspaper dist. by NEA, Inc.

So far we have talked about the ambiguity of individual terms or words. This is called *semantic ambiguity*. But sometimes we do not know which interpretation to give to a phrase or a sentence because its grammar or syntax admits of more than one interpretation. This is called *syntactic ambiguity* or *amphiboly*. Thus, if we talk about *the conquest of the Persians*, we might be referring either to the Persians' conquering someone or to someone's conquering the Persians. Sometimes the grammar of a sentence leaves open a great many possible interpretations. For example, consider the following sentence (from Paul Benacerraf):

Only sons marry only daughters.

One thing this might mean is that a person who is a male only child will marry a person who is a female only child. Again, it might mean that sons are the only persons who marry daughters and do not marry anyone else. Other interpretations are possible as well.

The process of rewriting a sentence so that one of its possible meanings becomes clear is called *disambiguating* the sentence. One way of disambiguating a sentence is to rewrite it as a whole, spelling things out in detail. That is how we disambiguated the sentence "Only sons marry only daughters." Another procedure is to continue the sentence in a way that supplies a context that forces one interpretation over others. Consider the sentence "Mary had a little lamb." Notice how the meaning changes completely under the following continuations:

1. Mary had a little lamb; it followed her to school.
2. Mary had a little lamb and then some broccoli.

Just in passing, it is not altogether obvious how we should describe the ambiguity in the sentence "Mary had a little lamb." The most obvious suggestion is that the word "had" is ambiguous, meaning "owned" on the first reading and "ate" on the second reading. Notice, however, that this also forces alternative readings for the expression "a little lamb." Presumably, it was a small, whole, live lamb that followed Mary to school, whereas it would have been a small amount of cooked lamb that she ate. So if we try to locate the ambiguity in particular words, we must say that not only the word "had" but also the word "lamb" are being used ambiguously. This is a reasonable approach, but another is available. In everyday speech, we often leave things out. Thus, instead of saying "Mary had a little *portion of meat derived from a lamb to eat*," we simply say "Mary had a little lamb," dropping out the italicized words on the assumption that they will be understood. In most contexts, such deletions cause no misunderstanding. But sometimes deletions are misunderstood, and this can produce ambiguity.

**EXERCISE I**

Show that each of the following sentences admits of at least two interpretations by (1) rewriting the sentence as a whole in two different ways and (2) expanding the sentence in two different ways to clarify the context:

Example: Kenneth let us down.

Rewriting: Kenneth lowered us.

Kenneth disappointed us.

Expanding: Kenneth let us down with a rope.

Kenneth let us down just when we needed him.

1. Barry Bonds (the baseball player) was safe at home.
2. I don't know what state Meredith is in.
3. Where did you get bitten?
4. The president sent her congratulations.
5. Visiting professors can be boring.
6. Wendy ran a marathon.
7. The meaning of the term "altering" is changing.
8. I don't want to get too close to him.
9. I often have my friends for dinner.
10. Slow Children Playing. (on a street sign)
11. Save Soap and Waste Paper. (on a sign during World War II)
12. In his will, he left \$1,000 to his two sons, Jim and John.
13. There is some explanation for everything.
14. She is an Asian historian.
15. Nobody may be in the lounge this evening.
16. Nobody came to the concert at 8 pm.

**EXERCISE II**

Follow the same instructions for the following actual newspaper headlines, many of which come from Columbia Journalism Review, editors, *Squad Helps Dog Bite Victim and Other Flubs from the Nation's Press* (Garden City, NY: Doubleday, 1980).

1. Milk Drinkers Turn to Powder
2. Anti-busing Rider Killed by Senate
3. Mrs. Gandhi Stoned in Rally in India
4. College Graduates Blind Senior Citizen
5. Jumping Bean Prices Affect the Poor
6. Tuna Biting off Washington Coast

7. Time for Football and Meatball Stew
8. Police Kill Man with Ax
9. Squad Helps Dog Bite Victim
10. Child Teaching Expert to Speak
11. Prostitutes Appeal to Pope
12. Legalized Outhouses Aired by Legislature
13. Police Can't Stop Gambling
14. Judge Permits Club to Continue Sex Bar
15. Greeks Fine Hookers
16. Survivor of Siamese Twins Joins Parents
17. Caribbean Islands Drift to the Left
18. Teenage Prostitution Problem Is Mounting
19. Miners Refuse to Work After Death
20. Police Begin Campaign to Run Down Jaywalkers
21. Red Tape Holds Up New Bridges
22. Juvenile Court to Try Shooting Defendant
23. Kids Make Nutritious Snacks
24. Study of Obesity Looks for Larger Test Group
25. Hospitals Sued by Seven Foot Doctors
26. Local High School Dropouts Cut in Half
27. Iraqi Head Seeks Arms
28. Drunk Gets Nine Months in Violin Case
29. Teacher Strikes Idle Kids
30. British Left Waffles on Falkland Islands
31. Stolen Painting Found by Tree
32. New Vaccine May Contain Rabies

### EXERCISE III

Poetry, songs, and jokes often intentionally exploit multiple meanings for effect. Find examples in poems, songs, and jokes that you like. Are these examples of ambiguity on the above definition? Why or why not?

## EQUIVOCATION

Ambiguity can cause a variety of problems for arguments. Often it produces hilarious or embarrassing side effects, and it is hard to get your arguments taken seriously if your listeners are giggling over an unintended double entendre in which one of the double meanings has risqué connotations.

Ambiguity can also generate bad arguments that involve the *fallacy of equivocation*. An argument is said to commit this fallacy when it uses the same expression in different senses in different parts of the argument, and this ruins the argument. Here is a silly example (from Carl Wolf):

Six is an odd number of legs for a horse.  
Odd numbers cannot be divided by two.  
 ∴ Six cannot be divided by two.

Clearly, “odd” means “unusual” in the first premise, but it means “not even” in the second premise. Consequently, both premises are true, even though the conclusion is false, so the argument is not valid.

Let’s consider another, more serious, example. In *Utilitarianism* (1861), John Stuart Mill claims to “prove” that “happiness is a good” with the following argument:

The only proof capable of being given that an object is visible is that people actually see it. The only proof that a sound is audible is that people hear it. In like manner the sole evidence it is possible to produce that anything is desirable is that people actually desire it. . . . [E]ach person, so far as he believes it to be attainable, desires his own happiness. This, however, being a fact, we have not only all the proof which the case admits of, but all which it is possible to require, that happiness is a good.<sup>1</sup>

Mill has sometimes been charged with committing a transparent fallacy in this passage. Specifically, the following argument is attributed to him:

- (1) If something is desired, then it is desirable.
  - (2) If it is desirable, then it is good.
- ∴ (3) If something is desired, then it is good.

Mill never presents his argument in this form, and it may be uncharitable to attribute it to him. Still, whether or not it is Mill’s way of arguing, it provides a good specimen of a fallacy of equivocation.

The objection to this argument is that the word “desirable” is used in different senses in the two premises. Specifically, in the first premise, it is used to mean “capable of being desired,” whereas in the second premise, it is used to mean “worthy of being desired.” If so, the argument really amounts to this:

- (1\*) If something is desired, then it is capable of being desired.
  - (2\*) If something is worthy of being desired, then it is good.
- ∴ (3) If something is desired, then it is good.

This argument is clearly not valid. To make the charge of equivocation stick, however, it has to be shown that the argument is not valid when the meaning

of the word “desirable” is used in the same sense in the two premises. This produces two cases to be examined:

- (1\*) If something is desired, then it is capable of being desired.
- (2\*\*) If something is capable of being desired, then it is good.
- .:(3) If something is desired, then it is good.

We now have a valid argument, but the second premise is not true, for sometimes people are capable of desiring things that are not good. The second way of restoring validity takes the following form:

- (1\*\*) If something is desired, then it is worthy of being desired.
- (2\*) If something is worthy of being desired, then it is good.
- .:(3) If something is desired, then it is good.

Again, we have a valid argument, but this time the first premise is false, since sometimes people do desire things that they should not desire. Thus, in both cases, altering the premises to produce a valid argument produces a false premise, so the argument cannot be sound.

This is a pattern that emerges when dealing with arguments that involve the fallacy of equivocation. When the premises are interpreted in a way that produces a valid argument, then at least one of the premises is false. When the premises are interpreted in a way that makes them true, then the argument is not valid. Here, then, is the strategy for dealing with arguments that may involve a fallacy of equivocation:

1. Distinguish the possible meanings of the potentially ambiguous expressions in the argument.
2. For each possible meaning, restate the argument so that each expression clearly has the same meaning in all of the premises and the conclusion.
3. Evaluate the resulting arguments separately.

If the argument fails whenever each term has a consistent meaning throughout the argument, then the argument is guilty of equivocation.

#### EXERCISE IV

Each of the following arguments trades on an ambiguity. For each, locate the ambiguity by showing that one or more of the statements can be interpreted in different ways.

1. We shouldn't hire Peter, because our company has a policy against hiring drug users, and I saw Peter take aspirin, which is a drug.
2. Man is the only rational animal, and no woman is a man, so women are not rational.

*(continued)*

3. My doctor has been practicing medicine for thirty years, and practice makes perfect, so my doctor must be nearly perfect.
4. Our cereal is all natural, for there is obviously nothing supernatural about it.
5. Ice cream is never all natural, since it never appears in nature without human intervention.
6. I have a right to spend all my money on lottery tickets. Therefore, when I spend all my money on lottery tickets, I am doing the right thing.
7. You passed no one on the road; therefore, you walked faster than no one.
8. Everything must have some cause; therefore, something must be the cause of everything.
9. The apostles were twelve. Matthew was an apostle. Hence, Matthew was twelve. (attributed to Bertrand Russell)
10. If I have only one friend, then I cannot say that I have any number of friends. So one is not any number. (from Timothy Duggan)
11. "Our bread does have fiber, because it contains wood pulp." (The Federal Trade Commission actually ordered the Continental Baking Company to indicate in their advertising that this is the kind of fiber in their Fresh Horizons bread.)
12. Anyone who tries to violate a law, even if the attempt fails, should be punished. People who try to fly are trying to violate the law of gravity. So they should be punished. (This argument is reported to have been used in an actual legal case during the nineteenth century, but compare Stephen Colbert, "Physics is the ultimate Big Government interference—universal laws meant to constrain us at every turn. . . . Hey, is it wrong that I sometimes want to act without having to deal with an equal and opposite reaction?"<sup>2</sup>)

#### DISCUSSION QUESTIONS

1. When a newspaper was criticized as a scandalous rumormonger, its editor responded with the following argument (as paraphrased by Deni Elliot). Does the editor's argument commit the fallacy of equivocation?  

It's not wrong for newspapers to pass on rumors about sex scandals. Newspapers have a duty to print stories that are in the public interest, and the public clearly has a great interest in rumors about sex scandals, since, when newspapers print such stories, their circulation increases, and they receive a large number of letters.
2. In the following passage, Tom Hill Jr. claims that a common argument against affirmative action commits a fallacy of equivocation. Do you agree that this argument equivocates? Why or why not?

Some think that the injustice of all affirmative action programs is obvious or easily demonstrated. [One argument] goes this way: "Affirmative action, by definition, gives preferential treatment to minorities and women. This is discrimination in their favor and against non-minority males. All discrimination by public institutions is unjust, no matter whether it is the old kind or the newer 'reverse discrimination.' So all affirmative action programs in public institutions are unjust."

This deceptively simple argument, of course, trades on an ambiguity. In one sense, to "discriminate" means to "make a distinction," to pay attention to a difference. In this evaluatively neutral sense, of course, affirmative action programs do discriminate. But public institutions must, and justifiably do, "discriminate" in this sense, for example, between citizens and noncitizens, freshmen and seniors, the talented and the retarded, and those who pay their bills and those who do not. Whether it is unjust to note and make use of a certain distinction in a given context depends upon many factors: the nature of the institution, the relevant rights of the parties involved, the purposes and effects of making that distinction, and so on.

All this would be obvious except for the fact that the word "discrimination" is also used in a pejorative sense, meaning (roughly) "making use of a distinction in an unjust or illegitimate way." To discriminate in this sense is obviously wrong, but now it remains an open question whether the use of gender and race distinctions in affirmative action programs is really "discrimination" in this sense. The simplistic argument uses the evaluatively neutral sense of "discrimination" to show that affirmative action discriminates; it then shifts to the pejorative sense when it asserts that discrimination is always wrong. Although one may, in the end, conclude that all public use of racial and gender distinctions is unjust, to do so requires more of an argument than the simple one (just given) that merely exploits an ambiguity of the word "discrimination."<sup>3</sup>

3. Many people argue that homosexuality is immoral because it is unnatural. Defenders criticize this argument for equivocating on various meanings of the term "unnatural." Distinguish some meanings of "unnatural." For each meaning, ask: Is homosexuality unnatural in that sense? Are acts immoral whenever (and because) they are unnatural in that sense? Why or why not?

## DEFINITIONS

It is sometimes suggested that a great many disputes could be avoided if people simply took the precaution of defining their terms. To some extent this is true. People do sometimes seem to disagree just because they are using terms in different ways, even though they agree on the nonverbal issues.

Nonetheless, definitions will not solve all problems, and a mindless insistence on definitions can turn a serious discussion into a semantic quibble. If you insist on defining every term, you will never be satisfied, because every definition will introduce new terms to be defined. Furthermore,

definitions themselves can be confusing or obfuscating as, for example, when an economist tells us:

I define “inflation” as too much money chasing too few goods.

Not only is this definition metaphorical and obscure, it also has a theory of the causes of inflation built into it.

To use definitions correctly, we must realize that they come in various forms and serve various purposes. There are at least five kinds of definitions that need to be distinguished:

1. *Lexical or dictionary definitions* are the most common kind of definition. We consult a dictionary when we are ignorant about the meaning of a word in a particular language. If you do not happen to know what the words “jejune,” “ketone,” or “Kreis” mean, then you can look these words up in an English, a scientific, and a German dictionary, respectively.

Except for an occasional diagram, dictionaries explain the meaning of a word by using other words that the reader presumably already understands. These explanations often run in a circle, such as when the *Oxford American Dictionary* defines “car” as “automobile” and “automobile” as “car.” Circular definitions can still be useful, because if you know what one of the terms in the circle means, you can use that background knowledge plus the definition to figure out what the other terms mean.

The goal of dictionary definitions is to supply us with factual information about the standard meanings of words in a particular language. As dictionary definitions are, in effect, factual claims about how people in general actually use certain words, dictionary definitions can be either accurate or inaccurate. The *Oxford American Dictionary* defines one meaning of “fan” as “a device waved in the hand or operated mechanically to create a current of air.” This is, strictly speaking, incorrect because a bellows also meets these conditions but is not a fan. Dictionary definitions can be criticized or defended on the basis of a speaker’s sense of the language or, more formally, by empirical surveys of what speakers accept as appropriate or reject as inappropriate uses of the term.

2. *Disambiguating definitions* specify a sense in which a word or phrase is or might be being used by a particular speaker on a particular occasion. (“When I said that the banks were collapsing, I meant river banks, not financial institutions.”) Disambiguating definitions can tell us which dictionary definition actually is intended in a particular context, or they can distinguish several meanings that might be intended. They can also be used to remove syntactic ambiguity or amphiboly. (“When I said that all of my friends are not students, I meant that not all of them are students, not that none of them are students.”)

Whether the ambiguity is semantic or syntactic, the goal of a disambiguating definition is to capture what the speaker intended, so such definitions can be justified by asking the speaker what he or she meant. This is a different

question than asking what a word means. Whereas dictionary definitions say what words mean or how they are used by most speakers of the language, a disambiguating definition focuses on a particular speaker and specifies which meaning that speaker intended on a particular occasion.

Such disambiguating definitions can be used in response to arguments that seem to commit the fallacy of equivocation. A critic can use disambiguating definitions to distinguish possible meanings and then ask, "Did you mean this or that?" The person who gave the argument can answer by picking one of these alternatives or by providing another disambiguating definition to specify what was meant. Speakers are sometimes not sure which meaning they intended, and then the critic needs to show that the argument cannot work if a single disambiguating definition is followed throughout. Whether one sides with the arguer or the critic, arguments that use terms ambiguously cannot be evaluated thoroughly without the help of disambiguating definitions.

3. *Stipulative definitions* are used to assign a meaning to a new (usually technical) term or to assign a new or special meaning to a familiar term. They have the following general form: "By such and such expression I (or we) will mean so and so." Thus, mathematicians introduced the new term "googol" to stand for the number expressed by 1 followed by one hundred 0s. Physicists use words like "charm," "color," and "strangeness" to stand for certain features of subatomic particles. Stipulative definitions do not report what a word means; they give a new word a meaning or an old word a new meaning.

Notice that if I say, "I stipulate that . . ." I thereby stipulate that . . . ; so such utterances are explicit performatives, and stipulation is a speech act. (See Chapter 2.) This explains why stipulative definitions cannot be false, since no performatives can be false. Stipulative definitions can, however, be criticized in other ways. They can be vague or ambiguous. They can be useless or confusing. Someone who stipulates a meaning for a term might go on to use the term with a different meaning (just as people sometimes fail to keep their promises). Still, stipulative definitions cannot be false by virtue of failing to correspond to the real meaning of a word, because they give that meaning to that word.

4. *Precising definitions* are used to resolve vagueness. They are used to draw a sharp (or sharper) boundary around the things to which a term refers, when this collection has a fuzzy or indeterminate boundary in ordinary usage. For example, it is not important for most purposes to decide how big a population center must be in order to count as a city rather than as a town. We can deal with the borderline cases by using such phrases as "very small city" or "quite a large town." It will not make much difference which phrase we use on most occasions. Yet it is not hard to imagine a situation in which it might make a difference whether a center of population is a city or not. As a city, it might be eligible for development funds that are not available to towns. Here a precising definition—a definition that draws a sharp boundary where none formerly existed—would be useful.

Precising definitions are, in effect, combinations of stipulative definitions and dictionary definitions. Like stipulative definitions, they involve a choice. One could define a city as any population center with more than fifty thousand people, or one could decide to decrease the minimum to thirty thousand people. Precising definitions are not completely arbitrary, however, because they usually should conform to the generally accepted meaning of a term. It would be unreasonable to define a city as any population center with more than seventeen people. Dictionary definitions, thus, set limits to precising definitions.

Precising definitions are also not arbitrary in another way: There can be good reasons to prefer one precising definition over another, when adopting the preferred definition will have better effects than the alternative. If development funds are to be distributed only to cities, then to define cities as having more than fifty thousand people will deny those funds to smaller population centers with, say, ten thousand people. Consequently, we need some reason to resolve the vagueness of the term "city" in one way rather than another. In this case, the choice might be based on the amount of funds available for development. In a more dramatic example, a precising definition of "death" might be used to resolve controversial issues about euthanasia—about what doctors may or must do to patients who are near death—and then our choices between possible precising definitions might be based on our deepest value commitments. In any case, we need some argument to show that one precising definition is better than other alternatives.

Such arguments often leave some leeway. Even if one can justify defining cities as having a minimum of fifty thousand people instead of ten thousand, one's reason is not likely to justify a cutoff at fifty thousand as opposed to forty-nine thousand. A different kind of defense would be needed if someone used a slippery-slope argument to show that it is unfair to provide development funds to one city with fifty thousand people but to deny such funds to its neighbor with only forty-nine thousand people. Against this kind of charge, the only way to defend a precising definition might be to show that some precising definition is needed, the cutoff should lie inside a certain general area, one's preferred definition does lie within that area, and no alternative is any better. Such responses might also apply to nearby alternatives, but they are still sometimes enough to support a precising definition. If responses like these are not available, then a precising definition can be criticized as unjustified.

5. *Systematic or theoretical definitions* are introduced to give a systematic order or structure to a subject matter. For example, in geometry, every term must be either a primitive (undefined) term or a term defined by means of these primitive terms. Thus, if we take points and distances as primitives, we can define a straight line as the shortest distance between two points. Then, assuming some more concepts, we can define a triangle as a closed figure with exactly three straight lines as sides. By a series of such definitions, the terms in geometry are placed in systematic relationships with one another.

In a similar way, we might try to represent family relationships using only the primitive notions of parent, male, and female. We could then construct definitions of the following kind:

"*A* is the brother of *B*." = "*A* and *B* have the same parents and *A* is male."

"*A* is *B*'s grandmother." = "*A* is a parent of a parent of *B* and *A* is female."<sup>4</sup>

Things become more complicated when we try to define such notions as "second cousin once removed" or "stepfather." Yet, by extending some basic definitions from simple to more complicated cases, all family relationships can be given a systematic presentation.

Formulating systematic definitions for family relationships is relatively easy, but similar activities in science, mathematics, and other fields can demand genius. It often takes deep insight into a subject to see which concepts are genuinely fundamental and which are secondary and derivative. When Sir Isaac Newton defined force in terms of mass and acceleration, he was not simply stating how he proposed to use certain words; he was introducing a fundamental conceptual relationship that improved our understanding of the physical world.

Such theoretical definitions can be evaluated on the basis of whether they really do help us formulate better theories and understand the world. Evaluating theoretical definitions often requires a great deal of empirical investigation. When water was defined as H<sub>2</sub>O,<sup>5</sup> this made it possible to formulate more precise laws about how water interacted with other chemicals. Other alternatives were available. Whereas molecules count as H<sub>2</sub>O, and hence as water, even if they contain unusual isotopes of hydrogen and oxygen, chemists could define water so that it would have to contain only the most common isotopes of hydrogen and oxygen. Why don't they? Because they discovered that differences among isotopes generally do not affect how molecules of H<sub>2</sub>O react with other chemicals. As a result, the simplest and most useful generalizations about the properties of water can be formulated in terms of H<sub>2</sub>O without regard to certain isotopes of hydrogen and oxygen. This illustrates one way in which choosing one theoretical definition over another can lead to a better theory.

Definitions can play important roles in the presentation of arguments, but demands for definitions can also hinder the progress of an argument. In the middle of discussions people often ask for definitions or even state, usually with an air of triumph, that everything depends on the way you define your terms. We saw in Chapter 2 that definitions are not always needed, and most issues do not turn on the way in which words are defined. When asked for a definition, it is appropriate to reply: "What sort of definition do you want, and why do you want it?" Of course, if you are using a word in a way that departs from customary usage, or using it in some special way of your own, or using a word that is too vague for the given context, or using a word in an ambiguous way, then the request for a definition is perfectly in order. In such cases, the demand for a definition represents an important move within the argument rather than a distraction from it.

**EXERCISE V**

Look up lexical or dictionary definitions for the following words. (For fun, you might try to guess the meanings of these words before you look them up, as in the game “Balderdash.”)

1. jejune
2. ketone
3. fluvial
4. xebec
5. plangent

**EXERCISE VI**

1. Give a stipulative definition for the word “klurg.”
2. Stipulate a word to stand for the chunks of ice that form under car fenders in winter.
3. Describe something that does not have a common name, for which it would be useful to stipulate a name. Explain how the name would be useful.

**EXERCISE VII**

Give precising definitions for the following words. In each case, supply a context that gives your precising definition a point.

1. book
2. alcoholic beverage
3. crime
4. warm
5. fast

**EXERCISE VIII**

Give disambiguating definitions for the following words. In each case, supply a context in which your definition might be needed to avoid confusion.

1. run
2. pen
3. game
4. painting
5. fast

**EXERCISE IX**

Using the notions of parents, male, and female as basic, give systematic definitions of the following family relationships:

1. *A* and *B* are sisters.
2. *A* and *B* are siblings.
3. *A* is *B*'s half-brother.
4. *A* is *B*'s niece.
5. *A* is *B*'s cousin.

**DISCUSSION QUESTION**

The U.S. federal criminal prohibition against torture (18 U.S.C. §§ 2340-2340A) prohibits conduct “specifically intended to inflict severe physical or mental pain or suffering.” On August 1, 2002, the U.S. attorney general’s office issued a statement that “severe” pain under the statute was limited to pain “equivalent in intensity to the pain accompanying serious physical injury, such as organ failure, impairment of bodily function, or even death.” (This interpretation was withdrawn in 2004.) What kind of a definition is this? Is it justified or not? What does this controversy show about the nature and importance of definitions?

**NOTES**

<sup>1</sup> John Stuart Mill, *Utilitarianism*, ed. George Sher (Indianapolis, IN: Hackett Publishing Company, 2001), 35.

<sup>2</sup> Stephen Colbert, *I Am America (And So Can You!)* (New York: Grand Central Publishing, 2007), 201.

<sup>3</sup> Thomas E. Hill Jr., “The Message of Affirmative Action,” in *Autonomy and Self-Respect*, ed. Thomas E. Hill Jr. (New York: Cambridge University Press, 1991), 193–94.

<sup>4</sup> Notice that in these definitions an individual word is not defined in isolation. Instead, a whole sentence containing the word is replaced by another whole sentence in which the defined word does not appear. Definitions of this kind are called “contextual definitions” because a context containing the word is the unit of definition. Dictionary, disambiguating, stipulative, and precising definitions can also be presented in this contextual form.

<sup>5</sup> If you doubt that the identity “Water is H<sub>2</sub>O” is used as a definition, just consider how you would react to someone who claims to have discovered some water that is *not* H<sub>2</sub>O. We would dismiss this person as linguistically confused, as the discovered stuff cannot properly be called “water” if it is not H<sub>2</sub>O.



## FALLACIES OF RELEVANCE

*This chapter will consider a different kind of defect in arguments. Fallacies of relevance arise when a premise, true or not, is not adequately related to the conclusion. Such irrelevance comes in endless varieties, but we will focus on two of the most common forms: arguments ad hominem and appeals to authority. Arguments of these kinds are not always fallacious, so we will discuss various factors that determine when such arguments are defective and when they are not.*

### RELEVANCE

---

In a good argument, we present statements that are true in order to offer a reason for some conclusion. One way to depart from this ideal is to state things that are true themselves, but have no bearing on the truth of the conclusion.

We might wonder why irrelevant remarks can have any influence at all. The answer is that we generally assume that a person's remarks are relevant, for this is one of the conditions for smooth and successful conversation (as Grice pointed out in his rule of Relevance, discussed in Chapter 2). That it is possible to exploit people by violating this natural assumption is shown in the following passage from *The Catcher in the Rye*.

The new elevator boy was sort of on the stupid side. I told him, in this very casual voice, to take me up the Dicksteins'....

He had the elevator doors all shut and all, and was all set to take me up, and then he turned around and said, "They ain't in. They're at a party on the fourteenth floor."

"That's all right," I said. "I'm supposed to wait for them. I'm their nephew."

He gave me this sort of stupid, suspicious look. "You better wait in the lobby, fella," he said.

"I'd like to—I really would," I said. "But I have a bad leg. I have to hold it in a certain position. I think I'd better sit down in the chair outside their door."

He didn't know what the hell I was talking about, so all he said was "oh" and took me up. Not bad, boy. It's funny. All you have to do is say something nobody understands and they'll do practically anything you want them to.<sup>1</sup>

It is clear what is going on here. When you offer something as a reason, it is conversationally implied that there is some connection between it and the thing you are arguing for. In most cases, the connection is obvious, and there is no need to spell it out. In other cases, the connection is not obvious, but in the spirit of cooperation others are willing to assume that the connection exists. In the present case, there seems to be no connection between having a bad leg and sitting in one particular chair. Why, then, does the elevator operator not challenge this statement? Part of the reason is that it is not easy to challenge what people say; among other things, it is not polite. But politeness does not seem to hold the elevator operator back; instead, he does not want to appear stupid. The person who offers a reason conversationally implies a connection, and we do not like to admit that we fail to see this connection. This combination of generosity and fear of looking stupid leads us to accept all sorts of irrelevant statements as reasons.

Fallacies of relevance are surprisingly common in everyday life. People often introduce irrelevant details or tangents in order to mislead by diverting attention from the real issue. The irrelevant distraction is sometimes described as a *red herring* (reportedly after a man who dragged a red herring across his trail in order to throw pursuing hounds off his scent). The best strategy for dealing with such tricks is simply to cross out all irrelevant claims and then see what is left. Sometimes nothing is left.

On the other hand, we should not be heavy-handed in making charges of irrelevance. Sometimes the occurrence of irrelevance is innocent; good arguments often contain irrelevant asides (as we saw in Chapter 5). More important, relevance is often secured by way of a conversational implication, so we really have to know what is going on in a given context to decide whether a remark is relevant or not. We can illustrate this last point by examining two kinds of arguments that often involve fallacies of irrelevance: *arguments ad hominem* and *appeals to authority*.

## AD HOMINEM ARGUMENTS

Literally, an argument *ad hominem* is an argument directed against a person who is making a claim rather than against that person's claim or argument for it. On the face of it, this move seems to involve irrelevance, for the character, social position, or status of a person should have nothing to do with the truth of what that person says or with the soundness or strength of that person's arguments. Even when protesters dress shabbily or fail to bathe, their clothing and hygiene show nothing about the legitimacy of their protest. A speaker's ethnicity, race, sex, or sexual orientation almost never give us any good reason to challenge the truth of what that person says or the soundness of his or her argument. And the fact that a judge was appointed by a liberal (or by a conservative) does not show that the judge's legal decisions

are incorrect or unfounded. Ad hominem fallacies very often deal in such irrelevant personal characteristics. They are often introduced just to distract us from the real point at issue.

In rare and unusual cases, however, a speaker's character or position is a reason to doubt the truth of what he says. Suppose that Lucy is suspected of committing murder, but Louie testifies that he was with her at the time of the murder. Then the prosecution shows that Louie provided a similar alibi for an accused murderer at ten trials in the past year, and every time he was found to have lied in exchange for money. Louie never testifies without being paid, he says whatever he is paid to say, and people do not hire him if they have any better defense. This background about Louie provides some reason to believe that what Louie said was false—that he was not with Lucy at the time of the murder. Lucy still might not have committed the murder, but we can't take Louie's word for it. Ad hominem arguments like this can be called *deniers*, since they deny the truth of what is said or the strength or soundness of an argument. Although most ad hominem deniers are fallacious, the case of Louie shows that a few are not.

A different kind of ad hominem argument questions a person's right to make a claim or present an argument. Imagine that the legislature is debating tax rates. During one session, Tad stands up and argues for a reduction in taxes. Tad can be criticized if he is not a legislator, because then he lacks the status that confers the right to speak in this setting. Even outside of any formal institution, if a neighbor tells someone that she ought to take her children to a certain church, the mother might respond, "Mind your own business, you busybody." Responses like this can be called *silencers*, because they revoke the right to speak without necessarily denying the truth of what is said.

A third variety of ad hominem argument is more subtle. Consider the following exchange:

**NORM:** The cold war is over, and bad relations between Cuba and the United States hurt both countries, so it is time for the United States to develop normal relations with Cuba.

**CLIFF:** Yeah, so you can make a bundle importing cigars from those commies.

Cliff's reply is not an attack on the truth of what Norm said. Nor is Cliff denying that Norm has a right to speak. Yet Cliff's remark is not without some relevance—it is not off the wall. Cliff is questioning Norm's motives. He suggests that Norm says what he says not because Norm believes it but only because Norm will make a lot of money if enough other people believe it. In a conversational exchange, we rely on the integrity of the person who is speaking, and when we have reasons to believe that the person's integrity is questionable, we sometimes say so. This is the significance of Cliff's remark. Cliff points to a fact that gives us some reason not to trust Norm's integrity in a discussion of U.S. relations with Cuba.

Cliff's attack might or might not be justified. If the only reason why Norm favors normal relations between the United States and Cuba is that this would enable Norm to make more money, then Cliff's ad hominem attack is well founded. But if Norm's real reason for saying what he does is that he honestly believes that normal relations would be beneficial both to the United States and to Cuba, then Norm's position does not depend on any lack of integrity. In that case, Cliff's attack is not well founded, even if it so happens that Norm would profit from normal relations.

Whether justified or not, ad hominem arguments of this third variety can be called *dismissers*, because they dismiss the speaker as untrustworthy and unreliable. Their point is not to deny the truth of the claim or the speaker's right to say it. Instead, a dismiser is supposed to show why the fact that this speaker supports a claim is not a good reason to believe that claim (or to deny it, for that matter).

These three variations are all ad hominem arguments because they start from premises about the person's character or status. Where they differ is in their conclusions: *Deniers* conclude that a claim is untrue or that an argument is unsound or weak. *Silencers* conclude that someone lacks the right to speak in a certain context. *Dismissers* conclude that someone is untrustworthy or unreliable. Each can be either justified or unjustified, so ad hominem come in six kinds that can be diagrammed like this:

Ad Hominem Arguments	Justified	Not Justified
<b>Deniers</b>	Louie, the hired perjurer	Shabby protesters
<b>Silencers</b>	Tad if he is not a legislator	Tad if he is a legislator
<b>Dismissers</b>	Cliff's reply if Norm lacks integrity	Cliff's reply if Norm does not lack integrity

What logicians usually call ad hominem fallacies are unjustified deniers. Even when the premises of such an argument are true, they are irrelevant to the conclusion. That makes them fallacies of relevance. Once you get used to spotting ad hominem fallacies, they seem common and obvious.

When assessing an ad hominem argument, the first step is to determine whether its conclusion is about someone's right to speak, about someone's reliability, or about the truth, soundness, or strength of what is claimed. The second step is to determine whether its premises provide adequate justification for its conclusion. These steps will enable you to place the argument in the above table, but they will often be neither easy nor obvious. Although perjurers for hire almost always lie, most people exhibit some middling degree of reliability. When people are known for passing on rumors without checking their truth, this might be a reason to doubt what they say when they pass on yet another rumor (even if it is not a reason to believe that what they say is false). In assessing what they say, it would be best to look for additional evidence. If none is available, then we need to ask how often their testimony is true on matters of this kind. Only by careful inspection of individual cases can we determine the strength of such ad hominem arguments.

## INCONSISTENCY

A related kind of argument occurs when someone is accused of inconsistency over time. If your neighbor says that the best time to prune roses is in the autumn, but the next spring she tells you that the spring is the best time to prune roses, then this inconsistency would give you some reason to doubt her expertise as a gardener. Maybe over the winter she got new information that changed her mind, but her reliability is in question until you have some explanation of why she would say different things at different times. If she wavers between contrary positions, then no more than half of her views can be correct, so you have reason to ask your neighbor, "Why have you changed your tune?" and "Why should I trust you now?"

What she is saying now still might be correct. Maybe spring is the right time to prune roses. Moreover, if she gave an argument for her claim (such as that pruned roses grow back more quickly in the spring, and it is better to prune roses when they grow back more quickly), then that argument still might be sound. Her current claim and argument do not depend on what she said last fall. For this reason, it is normally a fallacy to reject people's views on the basis of an inconsistency with their views at other times. Maybe they are right this time. Their current positions need to be assessed as they stand now. (Whether we want political leaders whose views blow with the wind is, of course, another issue.)

A different kind of inconsistency occurs in the traditional fallacy called *tu quoque*, a Latin term that means "you are another." When a parent tells a child to quit smoking, the child might respond, "Look who's talking. You've been smoking for years. If it's so bad, why don't you stop?" The force of this charge might be just that the parent is hypocritical or that one has no right to criticize others for doing something that one does oneself. If that is the point, then this response is a silencer, and it might or might not be justified. In any case, the parent's smoking does not give any reason at all to conclude that smoking is not bad. To use a *tu quoque* argument to reach *that* conclusion would be an unjustified denier (and an ad hominem fallacy). Even hypocrites can make true claims and give good arguments. Thus, to show that someone's claims and arguments are defective, one normally needs to look at those claims and arguments themselves, not at the behavior of the speaker.

## GENETIC FALLACIES

Instead of citing past beliefs or acts of a speaker, some ad hominem arguments aim at the source or origin of the speaker's belief. Stephen Colbert, for example, dismisses scientists by explaining how they got caught up in science: "They're physically awkward and lonely, so they spent their adolescence down by the creek studying the creatures that live there. 'I may be ridiculed at school,' they think, 'but a crawfish would never judge me.'"<sup>2</sup> This parody, of course, is supposed to show how silly it is to reject science because of its origin (even if this origin were plausible). When its origin is irrelevant to the truth of a claim, such arguments commit what is called the *genetic fallacy*.

Whole movements are sometimes accused of genetic fallacies. Marxists often reject opposing views because those views arose under capitalism. Freudians sometimes dismiss critics on the basis of how those critics were raised as children. Usually, the same kind of argument is available to their opponents as well. Critics of Freud sometimes cite his childhood and training to explain away his views. The fact that genetic arguments can be used just as well for contrary conclusions suggests that they do not really support either side.

The problem with genetic arguments is that lots of good ideas have questionable origins. Much mathematics originated within Pythagorean cults. Gravity and much of chemical theory were first discussed by alchemists who were trying to turn lead into gold. The structure of the benzene ring is reported to have come to Kekule von Stradonitz in a dream about a snake biting its tail. Early religions were close to magic and were used by rich and powerful leaders to control their subordinates. In all of these cases, the origin of the ideas cannot be used to refute those views or the arguments for them. Learning the genesis of an idea can help improve one's understanding of its content and of the process of discovery. Origins sometimes indicate where to look for evidence or for objections. Still, to evaluate an idea or an argument, one should focus on that idea or argument, not on its origin.

### EXERCISE I

For each of the following arguments, indicate whether it involves (1) an ad hominem denier against a speaker's claim or argument, (2) an ad hominem silencer against a person's right to speak, (3) an ad hominem dismiser against someone's trustworthiness, or (4) none of these. Explain your answer. Be sure to focus on what is explicitly said and not on what might be conversationally implied in each example.

Example: Sure, Sara says she saw me cheat in the game, but Sara's stupid, so you shouldn't pay any attention to her.

Answer: This is an ad hominem dismiser, since the point is that Sara is unreliable. The speaker does not deny what Sara says or that she has a right to say it.

1. Sure, Sadie says she saw me cheat, but it was very dark, and her vision is horrible, so she must have seen something else and thought it was me cheating.
2. Sure, Sam says he saw me cheat, but the only reason he says it is that he wants to win the game. He's a real jerk.
3. Sure, Steve says she saw me cheat, but he wasn't even playing the game. It's not his place to accuse those of us who were playing.
4. Sure, Sybill says she saw me cheat, but I didn't even take the exam, so I couldn't have cheated on it.
5. Sure, Sally says she saw me cheat, but she accuses everyone, and she's almost always wrong, so you should know that she is wrong this time, too.

**EXERCISE II**

Explain the point of each of the following remarks. Indicate whether each remark involves an ad hominem silencer, dismisser, denier, or none of these. Then say whether the argument provides an adequate justification for its conclusion, and why or why not.

1. The American Tobacco Company has argued for years that smoking is not really unhealthy, but what would you expect the company to say? It would take the same position regardless of any evidence, so I can't trust them.
2. The Joint Chiefs of Staff argue that the U.S. government needs to increase its military budget, but an opponent responds, "Well, of course, *they* will want as much money as they can get for their departments. They always ask for more money even though most of the time they don't really need it. So this time, again, they probably don't need it."
3. After Congress passes a military draft during a war, an opponent says, "If members of Congress were eligible for the draft, they would not vote for it. So we must not really need a draft."
4. Of course, the party in power is opposed to term limits. That's just because they want to stay in power.
5. The main opposition to tax reductions comes from people who depend on government programs funded by taxes, so they can hardly be impartial, but only those who are impartial should be allowed to speak on such a crucial issue for our whole country.
6. The main support for tax reductions comes from people who pay taxes, so their views can't be a reliable indicator of what the best policy is.
7. Very few citizens have studied the entire tax code, and nobody understands the effect of taxes on the economy, so we have little reason to believe them when they say that present tax policies will destroy the economy.
8. An economist cites recent trends in sales of raw materials as evidence of an upturn in the economy, and then a critic, who doubts the economist's prediction, responds, "If you're so smart, why ain't you rich?"
9. As a criticism of pro-choice activists, Ronald Reagan said, "I've noticed that everybody who is for abortion has already been born."
10. Attacking male opponents of abortion, a feminist claims, "Most opponents of abortion are men."
11. When a member of a fraternity argued for co-ed houses in place of fraternities, a critic responded, "When he quits his fraternity in protest and joins a co-ed house, then he will earn the right to criticize us."
12. When Fred argues at a fraternity meeting that his house should admit women, another member announces, "Let me remind you all that Fred held exactly the opposite position last year."
13. Let he who is without sin among you cast the first stone. (John 8:7)

**DISCUSSION QUESTIONS**

1. In a heated discussion, people will sometimes ask an opponent, "Why are you being so defensive?" This is obviously a rhetorical question. What is the point of this question? Does it implicitly involve an ad hominem fallacy?
2. In the biblical story of Job, Job is described as a person who "was blameless and upright, one who feared God and turned away from evil" (Job 1:1). Satan challenges God to allow him to subject Job to the worst calamities to see if Job's faith will remain unchanged. After the most extreme misfortunes, Job finally cries out and asks why he should be made to suffer so (Job 38:1–4):

Then the Lord answered Job out of the whirlwind:  
 Who is this that darkens counsel by words without knowledge?  
 Gird up your loins like a man.  
 I will question you, and you shall declare to me.  
 Where were you when I laid the foundation of the earth?  
 Tell me, if you have understanding.

Does God's response to Job involve an ad hominem silencer, dismisser, or denier? Is it justified? Why or why not?

**APPEALS TO AUTHORITY**

Often in the midst of an argument, we cite an authority to back up what we say. As we saw in Chapter 3, this is a standard way of offering assurances. In citing an authority, instead of giving reasons for what we say, we indicate that someone (the cited authority) could give them.

Although logicians sometimes speak of the *fallacy* of appealing to authorities, there is often nothing wrong with citing authorities or experts to support what we say. An authority is a person or institution with a privileged position concerning certain information. Through training, a doctor is an expert on certain diseases. A person who works in the Department of Agriculture can be an expert on America's soybean production. Someone who grew up in the swamps might be an expert on trapping muskrats. Because some people stand in a better position to know things than others, there is nothing improper about citing them as authorities. In fact, an appeal to experts and authorities is essential if we are to make up our minds on subjects outside our own range of competence.

At the same time, appeals to authority can be abused, and there are some obvious questions we should ask whenever such an appeal is made. Most obviously, we should always ask whether the person cited is, in fact, an authority at all. Moreover, it is not enough to be an authority in some area or other. We need to ask whether the person cited is an authority in the particular area under discussion. If the answer to this question is "No," then we are

dealing with a *fallacy of relevance*. For example, being a movie star does not qualify a person to speak on the merits of a particular brand of toothpaste. Endorsements by athletes of hair creams, deodorants, beer, and automobiles are in the same boat. Of course, we have to be careful in making this charge. It is possible that certain athletes make systematic studies of deodorants before giving one deodorant their endorsement. But it is not likely.

Most people realize that athletes, movie stars, and the like are featured in advertisements primarily to attract attention and not because they are experts concerning the products they are endorsing. It is more surprising how often the wrong authorities are brought in to judge serious matters. To cite one example, Uri Geller had little difficulty convincing a group of distinguished British scientists that he possessed psychic powers. In particular, he was able to convince them that he could bend spoons by mental powers alone. In contrast, James Randi, a professional magician, had little difficulty detecting and duplicating the tricks that bamboozled the scientific observers. The remarkable feature of this case was not that a group of scientists could be fooled by a magician, but rather that these scientists assumed that they had the expertise necessary to decide whether a paranormal phenomenon had taken place or not. After all, the most obvious explanation of Geller's feats was that he had somehow cheated. To test this possibility, what was needed was not a scientist with impeccable scholarly credentials, but a magician who could do the same tricks himself and therefore knew what to look for.<sup>3</sup>

It is, of course, difficult to decide whether someone is an expert in a field when you yourself are not, but certain clues will help you make this decision. If the supposed authority claims to have knowledge of things that he or she could not possibly possess (for example, about private conversations the person could not have heard), then you have little reason to trust other things that person has to say. You know that he or she has no qualms about making things up. Furthermore, it is often possible to spot-check certain claims in order to make sure that they are correct. It may take one expert to determine another, but it often takes little more than good common sense and an unwillingness to be fooled to detect a fraud.

Even when it is clear that the cited person is an expert in the appropriate field, we can still ask whether the question is of the kind that can now be settled by an appeal to experts. One sign that a question cannot yet be settled by experts is that experts in that area do not agree with each other. It does not do much good to cite one authority in support of a claim if another authority with just as much expertise would endorse the opposite claim. Moreover, even the best experts sometimes simply get things wrong. For example, in 1932 Albert Einstein, who was surely an expert in the field, declared, "There is not the slightest indication that [nuclear] energy will ever be obtainable. It would mean that the atom would have to be shattered at will." Just a year later, the atom was, in fact, split. Even so, a leading British physicist, Ernest Lord Rutherford, insisted that the splitting of the atom would not lead to the development of nuclear power, saying, "The energy produced by the atom is

a very poor kind of thing. Anyone who expects a source of power from the transformation of these atoms is talking moonshine.”<sup>4</sup> Given the knowledge available at the time, both Einstein and Rutherford may have been justified in their claims, but their assertions were, after all, more speculations than scientifically supported statements of fact. The lesson to be learned from this is that the best experts are sometimes fallible, and become more fallible when they go beyond established facts in their discipline to speculate about the future.

Although the next question may seem obvious, we often forget to ask whether the authority has been cited correctly. When a person cites an authority, he or she is making a factual claim that so-and-so holds some particular view. Sometimes the claim is false. If someone told you, “According to medical authorities, the rash from poison ivy is contagious when it is oozing,” you would probably believe it. In fact, the citation is incorrect. According to medical authorities, the rash from poison ivy is never contagious. Yet many people hold that it is contagious, and they think that they have medical opinion on their side. It is hard to deal with people who cite authorities incorrectly, for we do not carry an almanac or encyclopedia around with us. We can check it on the Internet if we have a computer with us, but that might seem impolite. Still, it is a good idea to spot-check appeals to authority, for people often twist authorities to support their own opinions.

It is also worth asking whether the cited authority can be trusted to tell the truth. To put this more bluntly, we should ask whether a particular authority has any good reason to lie or misrepresent facts. Presumably, the officials who know most about food production in China will be the heads of the various agricultural bureaus. But it would be utterly naive to take their reports at face value. Inadequate agricultural production has been a standing embarrassment of the Chinese economy. As a consequence, there is pressure at every level to make things look as good as possible. Even if the state officials were inclined to tell the truth, which is a charitable assumption, the information they receive is probably not very accurate.

Experts also lie because it can bring fame and professional advancement. Science, sometimes at the highest level, has been embarrassed by problems of the falsification and misrepresentation of data. Consider the case of Sir Cyril Burt’s research on the inheritance of intelligence. Burt wanted to show that there is a significant correlation between the IQs of parents and their children. The difficulty was to find a way to screen out other influences—for example, that of home environment. To overcome this, Burt undertook a systematic study of identical twins who had been separated at birth and raised in various social settings. His study revealed a very high correlation between the IQs of these twins, and that gave strong reason to believe that IQ, to some significant extent, depends on heredity rather than environment.

Unfortunately, Burt’s data, or at least a significantly large portion of them, were cooked—that is, made up. It is interesting that Burt’s bogus research could go unchallenged for so long. It is also interesting how he was finally unmasked. First, to many his results seemed too good to be true. He claimed

to have found more than fifty identical twins who had been separated at birth and raised in contrasting environments. Given the rarity of such creatures, that is a very large number to have found. Second, the correlations he claimed to find were extremely high—indeed, much higher than those usually found in research in this area. Both of these facts raised suspicions. Stephen Jay Gould describes Burt's final undoing as follows:

Princeton psychologist Leon Kamin first noted that, while Burt had increased his sample of twins from fewer than twenty to more than fifty in a series of publications, the average correlation between pairs for IQ remained unchanged to the third decimal place—a statistical situation so unlikely that it matches the vernacular definition of impossible. Then, in 1976, Oliver Gillie, medical correspondent of the London *Sunday Times*, elevated the charge from inexcusable carelessness to conscious fakery. Gillie discovered, among many other things, that Burt's two “collaborators” . . . the women who supposedly collected and processed his data, either never existed at all, or at least could not have been in contact with Burt while he wrote the papers bearing their names.<sup>5</sup>

Of course, Burt's claims still might be correct: genes and IQ might be correlated. Nonetheless the point here is just that Burt and his studies should not be trusted as authorities. Outright fraud of this kind by someone so prominent is rare, but even a few cases provides a reason for being suspicious of authorities, at least when their results have not been given independent confirmation.

One last question we can ask is why the appeal to authority is being made at all. To cite an authority is to give assurances. As we noticed in Chapter 3, people usually give assurances to strengthen weak points in their arguments. It is surprising how often we can see what is wrong with an argument just by noticing where it is backed by appeals to authority. Beyond this, we should be suspicious of arguments that rely on too many authorities. (We might call this the fallacy of *excessive footnotes*.) Good arguments tend to stand on their own.

To summarize, reliance on experts and authorities is unavoidable in our complicated and specialized world. Yet we still need to be critical of appeals to authority by asking these questions:

1. Is the cited authority in fact an authority in the appropriate area?
2. Is this the kind of question that can now be settled by expert consensus?
3. Has the authority been cited correctly?
4. Can the cited authority be trusted to tell the truth?
5. Why is an appeal to authority being made at all?

If the answers to questions 1–4 are “Yes,” then the appeal to authority is probably justified. Still, even the best authorities make mistakes, so the conclusion of any appeal to authority might turn out to be false. We can reduce errors by appealing to better authorities, but no authority can guarantee the truth.

**EXERCISE III**

Answer the five questions in the text about each of the following appeals to authority, and then decide whether each appeal to authority is legitimate or fallacious.

1. The surgeon general says that smoking is hazardous to your health, so it is.
2. The surgeon general says that abortion is immoral, so it is.
3. Michael Jordan says that Air Jordan sneakers are springier, so they must be springier.
4. This must be a great movie, because the billboard says that *Time* magazine called it “terrific.”
5. My friend Joe says that this new movie is hilarious, so it must be worth watching.
6. Ben and Jerry’s ice cream must be the best, because Fat Fred eats more ice cream than anyone else I know, and he says that Ben and Jerry’s is the best.
7. There must be life on other planets, because many great scientists are looking for it, so they must think it is there.
8. Lefty Lopez must be the best pitcher of the year, because he won the Cy Young Award (awarded by the Baseball Writers Association to the best pitcher of the year).
9. Vanna must be the most beautiful woman in America, because she won the Miss America contest.
10. There were 250,000 protesters at the rally, because its organizers gave that figure.
11. There were 25,000 protesters at the rally, because its opponents said so.
12. True Christians ought to give away all their money, because the Bible says, “Blessed are the poor.”

**MORE FALLACIES OF RELEVANCE**

Questions like those used to evaluate appeals to authority can also be used to assess some other common styles of reasoning that are often accused of being fallacious.

**APPEALS TO POPULAR OPINION**

Here is one example:

The American people are convinced that, if we get involved in North Korea, we will be stuck there for a long time. So we shouldn’t invade in the first place.

This argument, of course, depends on suppressed premises. On one reconstruction, the argument is that, because Americans fear getting stuck in North Korea, they oppose American involvement, and a democratic

government should not do what the people oppose, so we should not get involved in North Korea. The argument also seems to suggest another reason why America should not get involved—namely, that if we do, our troops will be stuck in North Korea. The only reason given for believing this is that lots of American people believe it. So the argument seems to be that, because so many Americans believe it, it must be true.

Such an argument is not an appeal to authority, since no person is claimed to be an authority or an expert. Instead, the argument is an *appeal to popular opinion*. When the popular opinion is supposed to have been shared for a long time, the argument can be called an *appeal to tradition*. Such arguments assume that, when many people agree on some issue or agree for a long time, they are likely to be right. This assumption is often incorrect. An opinion might be shared by many people just because they all learned it from a common source, such as television or some prominent politicians. Then the shared opinion is not reliable unless its source is reliable. Of course, the shared opinion *might* be true; America *might* get stuck in North Korea if it got involved. But the argument for this conclusion is still fallacious, because the mere fact that an opinion is widely held is not enough to show that the opinion is true.

Although such appeals to popular opinion are often fallacious, there are also some areas where popular opinion is evidence of truth. If most people who read a book think that the book is entertaining and easy to understand, then it is entertaining and easy to understand. If most people think that the sky looks blue, this is evidence that the sky does look blue. Thus, not all appeals to popular opinion are defective or fallacious.

To determine whether or not a particular appeal to popular opinion is fallacious, we need to ask questions that are much like the questions we asked about appeals to authority. These include:

1. Is this opinion actually widely held?
2. Is this the kind of area where popular opinion is likely to be right?
3. Why is an appeal to popular opinion being made at all?

Even when superficial examination reveals that an appeal to popular opinion is fallacious, such arguments still seem to convince many people. This might be because many people want to agree with others so that they will be popular and will not have to think for themselves.

## APPEALS TO EMOTION

Other appeals are not to beliefs but, instead, to emotions. One common form of appeal to emotion is an *appeal to pity*. Defense lawyers often dwell on the sad circumstances in which a defendant grew up or on how badly the defendant's family will be hurt if the defendant goes to prison. Such an appeal to pity might show that the defendant should not receive the maximum sentence. But when such an appeal to pity is used to argue that the

defendant is not guilty or should not be found guilty, then the argument is almost always fallacious.

*Appeals to fear* are also common, especially since al Qaeda's destruction of the World Trade Center towers in New York on September 11, 2001. After another terrorist attack occurred in London in 2005, television news anchors and talk-show hosts flashed headlines like "Who's at risk?", "Are we next in America?", "How safe are we in America?", "How prepared are we?", "Can we prevent a subway or a bus attack in the US?", and "You have to wonder, will we ever truly feel safe again?" To be fair, television stations asked these questions because they knew that their audiences wanted answers to those very questions. A problem arises only when the media stirs up fears and fails to provide a balanced estimation of the real dangers. Such fearmongering is parodied in Stephen Colbert's segment, "The Threat Down," which lists more and more outrageous fears. An even more serious problem arises when politicians use exaggerated fears to gain support for costly policies that could not be justified by the actual threats that people face. Of course, it is controversial how much fear and which policies are justified. Liberals accuse Conservatives of distortion when Conservatives use fear of terrorism to justify their counterterrorist policies; Conservatives accuse Liberals of exaggeration when Liberals paint a scary picture of the effects of global warming. Defenders of any particular policy can reply that it really is needed to ward off peril. Some fears are justified, including some fears of terrorism and of global warming, of course. Appeals to fear are fallacious only when they are overdone or exaggerated in order to lead people away from an accurate assessment of the risks that really exist. But this happens all too often; so, whenever anyone appeals to fear, we need to ask whether those fears are being amplified and abused.

Outrage is another emotion that many arguments appeal to. On Day 3 of the 2004 Republican National Convention, for example, Democratic Senator Zell Miller proclaimed, "Today's Democratic leaders see America as an occupier, not a liberator; and nothing makes this marine madder than someone calling American troops occupiers rather than liberators!" This line drew thunderous applause, even though Miller did not give any reason either against calling American troops occupiers or for describing them as liberators. Nonetheless, a receptive audience will tend to assume that, if such an impressive speaker is *that* outraged (and if others in the audience share that outrage), then there must be something terribly objectionable about whatever the outrage is directed against. To assess such appeals to outrage, as for other emotions, we need to become aware of these common assumptions so that they can be critically evaluated.

Appeals to emotion can also be positive. Many advertisements work by linking a product to positive feelings. Everyone knows that a car does not become better just because it is displayed in beautiful scenery, but it is amazing how much the scenery in advertisements can affect people's inclinations to buy a certain car. Similarly, advocates of a treaty or government program often paint pictures of how wonderful life will be if the treaty or program works

out well. Such appeals to emotion might provide some reason to adopt their plan, but these arguments can be very misleading if it is unlikely that everything will work out so well and if serious dangers will arise when something goes wrong. Thus, even when emotional reactions are relevant to some extent, one must be careful not to let them cloud the other side of the issue.

#### EXERCISE IV

For each of the following arguments, indicate whether it is an appeal to popular opinion, an appeal to tradition, or an appeal to emotion. (The argument might fit into more than one of these categories. If so, explain why.) Then determine whether it is fallacious, and why.

1. For centuries throughout Europe, women were burned for being witches, so there must have been lots of witches.
2. There must be life on other planets, because most people think there is. Just read a few tabloids.
3. Most people who live in the United States think that it is the greatest country ever, so it must be.
4. There are more Buddhists than followers of any other religion, so there must be more truth in Buddhism.
5. Incest must be immoral, because people all over the world for many centuries have seen it as immoral.
6. The Golden Rule is accepted in almost every system of ethics both in the past and in the present, so there is probably something to it.
7. Chris must not be guilty, because twelve jurors, who saw all the evidence, agreed on a verdict of not guilty.
8. “Polls show an overwhelming majority of the American people want a lot less immigration or even an immigration moratorium. . . . These are persistent results over time. Most of the people cannot be wrong all of the time!” (from an advertisement placed by Federation for American Immigration Reform, *Atlantic Monthly* [June 1995], 67).

#### NOTES

<sup>1</sup> J. D. Salinger, *The Catcher in the Rye* (New York: Bantam Books, 1951), 157–58.

<sup>2</sup> Stephen Colbert, *I Am America (And So Can You!)* (New York: Grand Central Publishing, 2007), 193.

<sup>3</sup> For an entertaining and instructive account of this case, see James Randi, *The Magic of Uri Geller* (New York: Ballantine Books, 1975).

<sup>4</sup> Both quotations are from Christopher Cerf and Victor Navasky, *The Experts Speak* (New York: Pantheon Books, 1984), 215. This work contains a marvelous collection of false and sometimes just plain stupid things that have been claimed by experts. One notable example is the remark made by the Union general John B. Sedgwick just before being fatally shot in the head by a Confederate marksman: “They couldn’t hit an elephant at this dist—” (135).

<sup>5</sup> Stephen Jay Gould, *The Mismeasure of Man* (New York: Norton, 1981), 235.



## FALLACIES OF VACUITY

*Arguments are vacuous when they don't go anywhere. This happens in two main ways. Sometimes an argument begins by assuming its conclusion, so the argument makes no real progress beyond its own assumptions. In other cases, the argument's conclusion is empty, so the argument has nowhere in particular to go. Both kinds of argument are fallacious and vacuous, so we call them fallacies of vacuity. Circular arguments and arguments that beg the question fall into this category. So do positions that make themselves immune to criticism by being self-sealing.*

## CIRCULARITY

---

One purpose of arguments is to establish the truth of a claim to someone who doubts it. In a typical situation, one person, *A*, makes a claim; another person, *B*, raises objections to it; then *A* tries to find arguments that respond to the objections and justify the original statement. Schematically:

- A* asserts that *p* is true.
- B* raises objections *x*, *y*, and *z* against it.
- A* then offers reasons to overcome these objections.

What must *A*'s responses be like to meet *B*'s objections? To start with the simplest case, *A* cannot meet *B*'s challenge simply by repeating the original assertion. If someone is maintaining that terrorists can't be stopped without torture, it will not help to offer as a justification for this the very claim that is in dispute—that terrorists can't be stopped without torture. The argument would then look like this:

- Terrorists can't be stopped without torture.
- ∴ Terrorists can't be stopped without torture.

This argument is, of course, *valid*, since the premise cannot be true without the conclusion being true as well. Furthermore, if the premise is true, then the argument is also *sound*. All the same, the argument has no force in this conversational setting because any objection that *B* has to the conclusion is straight off an objection to the premise, since they are identical.

Unfortunately, people usually do not make it so easy to tell when they reason in a circle. Often, circular reasoning is disguised by restating the conclusion in different words. Someone might argue that terrorists can't be stopped without torture, because, if you do not use torture, there is no other way to stop terrorists. This premise means the same as the conclusion, so this reasoning is still circular.

Another way to hide circularity is by suppressing the premise that repeats the conclusion. (See Chapter 5 on suppressed premises.) Suppose someone argues that terrorists cannot be stopped without torture, because they are so callous that their goal is to kill and maim innocent civilians. This argument depends on the suppressed premise that anyone whose goal is to kill and maim innocent civilians cannot be stopped without torture. If terrorists are then defined as people whose goal is to kill and maim innocent civilians, then this suppressed premise reduces to the conclusion that terrorists cannot be stopped without torture. So this argument is also circular.

Yet another trick is to put forward a statement first as a conclusion to be proved, and then only much later—after several subarguments or tangents—use the same statement as a premise on its own behalf. Consider this simple argument:

The only way to prevent terrorists from committing their horrible crimes is to inflict enough pain on them either to scare them off or to force them to reveal information that enables the police to head off terrorist attacks. Because these are the only methods that work, we cannot reason with them or talk them into giving up. We cannot make friends or sign a treaty with them. We cannot buy them off or satisfy their demands. Therefore, terrorists cannot be stopped without torture.

If the first sentence is supposed to provide a reason for the next three sentences, then those three sentences cannot later be used as a reason for the last sentence without the whole argument becoming circular, because the last sentence, "Terrorists can't be stopped without torture," means pretty much the same as the first sentence, "The only way to prevent terrorists . . . is to inflict enough pain on them. . ." Although this trick is often harder to detect in a long and complex argument, such reasoning is still indirectly circular if any premise in a chain of arguments repeats or restates the eventual conclusion. Thus, we have *circular reasoning* if and only if one of the premises that is used directly or indirectly to support a conclusion is equivalent to the conclusion itself.

## BEGGING THE QUESTION

Reasoning in a circle is normally bad reasoning, but it is not easy to say exactly what is bad about it. The problem with circular reasoning becomes clearer when we notice that the same basic defect is shared by arguments that are not strictly circular. Instead of arguing, "Terrorists can't be stopped

without torture, so they can't," we could avoid circularity by adding a few words to get this new argument:

If terrorists can be stopped without torture, then I'm a monkey's uncle.  
I'm not a monkey's uncle.

. . . Terrorists can't be stopped without torture.

This argument is not circular, because neither premise repeats the conclusion. It is also valid, and it might even be sound. Still, it is no good as an argument for the same reason as "Terrorists can't be stopped without torture, so it can't." The problem lies in its first premise. The first premise is a conditional, and its consequent ("I'm a monkey's uncle") is obviously false, so that first premise as a whole is false unless its antecedent ("Terrorists can be stopped without torture") is also false. However, that means that the conclusion must be true in order for the first premise to be true. Thus, one could not have any reason to believe the first premise if one did not already have the very same reason to believe the conclusion. In short, one cannot have any independent reason for the premise.

More generally, we can say that an argument *begs the question* in a context if and only if (1) it depends on a premise that is not supported by any reason that is independent of the conclusion, and (2) there is a need for such an independent reason.

To say that an argument begs the question in this sense is not just to say that it raises the question. That is what a sports announcer means, for example, when she says, "His injury begs the question of whether he will return in time for the playoffs." This common use of the phrase "begs the question" is separate from the fallacy, but they are not completely unrelated. An argument can also be seen as begging the question when its context raises the question of why anyone who denies its conclusion should accept its premises and when that question has no adequate answer.

More precisely, the need for an independent justification arises from the context and the purpose for which the argument is being used. A premise *needs* support from an independent reason, for example, when it is in dispute or subject to objection and the arguer's goal is to give an audience some reason to accept the premise and, on that basis, to accept the conclusion. That such a need for an independent reason exists but is not satisfied explains why the argument can be criticized by saying that it commits the fallacy of *begging the question*.

This fallacy is often very hard to detect, both because it is affected by the context and because there are many ways to hide the fact that a premise depends on the conclusion. Consequently, people often use arguments that beg the question when they have nothing better to say, especially on a controversial issue. It is common, for example, to hear an argument something like the following:

It's always wrong to murder human beings.

Capital punishment involves murdering human beings.

. . . Capital punishment is wrong.

Here the first premise is true by definition, since calling something “murder” implies that it is a wrongful killing. The second premise is, however, question begging, for calling capital punishment murder assumes the point at issue—that capital punishment is wrong. Capital punishment is obviously *killing*, but whether it is murder in the sense of *wrongful killing* is precisely what is disputed. As a result, anyone who objects to the conclusion would or should raise exactly the same objections to the second premise, and one could not give any adequate reason for the second premise without first arguing for the conclusion.

More subtly than this, opponents of abortion typically refer to the human fetus as an unborn baby or simply as a baby. It may seem a matter of indifference how the fetus is referred to, but this is not true. One of the central points in the debate over abortion is whether the fetus has the status of a person and, thus, has the rights that any person has. It is generally acknowledged in modern societies that babies are persons and therefore have the rights of persons. By referring to the fetus as an unborn baby (or simply as a baby), a point that demands argument is taken for granted without argument. That counts as begging the question. Of course, many opponents of abortion argue for the claim that a human fetus has the moral status of a person and, thus, do not beg this central question in the debate. Still, if they give no such independent argument, then they do beg the question.

Similarly, if someone argues for the pro-choice position simply on the grounds that a woman has a right to control the destiny of her own body, this also begs an important question, because it takes for granted the claim that the fetus is part of a woman’s body, not an independent being with rights of its own. Of course, defenders of the pro-choice position need not beg the question in this way, but they often do. Whether a particular argument or premise is question begging will depend on whether there is a need for an independent reason, which in turn depends on the context in which the argument is given. One way for an argument to beg the question is for it to rely, either explicitly or implicitly, on an unsupported premise that is a matter of dispute in the particular argumentative context. Thus, referring to a human fetus as a baby will be question begging in contexts in which the moral status of the fetus is at issue, but it may not be question begging when this is not an issue.

Because begging the question depends in this way on context, we should be careful before charging opponents with begging the question. Some people charge every opponent with begging the question, almost like a knee-jerk reaction. However, even if an opponent uses a premise that you reject, this does not yet show that the argument begs the question, since your opponent might have plenty of independent evidence for the premise. Before you accuse people of begging the question, you should ask them to give you their reasons for the disputed premise. If

they can come up with an independent reason, then they did not beg the question, and you might learn something from them. However, if they do not have any independent reason for the premise, then they did indeed beg the question.

### EXERCISE I

For each of the following arguments, does it involve circular reasoning? Does it beg the question in any context? If so, in which contexts? Explain your answers.

1. A student of mine told me that I am her favorite professor, and I know that she is telling the truth, because no student would lie to her favorite professor.
2. Intoxicating beverages should be banned, because they can make people drunk.
3. Capitalism is the only correct economic system, because without it free enterprise would be impossible.
4. Free trade is good for the country, because it brings the country all of the advantages of an unimpeded flow of goods.
5. Gun-control laws are wrong, because they violate the citizen's right to bear arms.
6. When *B* applies for a job from *A*:  
**A:** How can we know that you are trustworthy?  
**B:** Mr. Davidson will write me a recommendation.  
**A:** But why should we trust him?  
**B:** I assure you that he is honest and accurate.
7. The Bible is the inerrant word of God, because God speaks only the truth, and repeatedly in the Bible God tells us that the Bible consists of His words.
8. We have to accept change, because without change there is no progress.
9. Premarital sex is wrong, because premarital sex is fornication, and fornication is a sin.
10. The drinking age should be lowered to eighteen, because eighteen-year-olds are mature enough to drink.
11. We should never give security clearances to homosexuals, because they can be blackmailed into revealing classified information. They are subject to blackmail, because we will revoke their security clearances if we find out they are gay.
12. People with suicidal tendencies are insane, because they want to kill themselves.
13. Jeffrey can't really be insane, because he says he is.

**DISCUSSION QUESTIONS**

1. Explanations are often presented in the form of arguments that sometimes seem circular. Are the following arguments circular? Do they beg the question? Are they defective in some other way? Why or why not? More generally, when, if ever, can circular arguments provide good explanations?

A. **TOM:** Why are so many people moving out of Claremont this year?

**SUE:** Because its economy is going down so fast.

**TOM:** But why is its economy going down so fast?

**SUE:** Because so many people are moving out of town.

B. **AMY:** Why is Jarred going down on the seesaw right now?

**JOHN:** Because Jeremiah is going up on the other side of the seesaw.

**AMY:** But why is Jeremiah going up right now?

**JOHN:** Because Jarred is going down.

2. Explain John Stuart Mill's argument in the following passage (from *A System of Logic* [London, 1843], book 2, chapter 3, section 2). Do you agree? Why or why not?

It must be granted that in every syllogism, considered as an argument to prove the conclusion, there is a *petitio principii* [a begging of the question]. When we say,

All men are mortal.

Socrates is a man.

∴ Socrates is mortal;

it is unanswerably urged by the adversaries of the syllogistic theory that the proposition, "Socrates is mortal," is presupposed in the more general assumption, "All men are mortal"; that we cannot be assured of the mortality of all men unless we are already certain of the mortality of every individual man. . . . That, in short, no reasoning from generals to particulars can, as such, prove anything, since from a general principle we cannot infer any particulars but those which the principle itself assumes as known.<sup>1</sup>

**SELF-SEALERS**

It is characteristic of certain positions that no evidence can *possibly* refute them. This may seem to be a wonderful feature for a position to have. In fact, however, it *usually* makes the position useless. We can start with a silly example. A Perfect Sage claims to be able to predict the future in detail. The Perfect Sage's predictions take the following form:

Two weeks from today at 4:37 you are going to be doing *exactly* what you will be doing.

Of course, whatever you are doing at that time will be exactly what you are doing, so this prediction cannot possibly be wrong. But this is only because it does not tell us anything in particular about the future. *Whatever* happens,

the prediction is going to be true, and this is just what is wrong with it. The prediction is *empty* or *vacuous*.

People do not, of course, go around making predictions of this kind, but they do sometimes hold positions that are empty or vacuous in much the same way. A clairvoyant claims to be able to predict the future, but every time a prediction fails, she says that this just proves that someone set up bad vibrations that interfered with her visions. So, if the prediction turns out to be true, she claims that this shows her clairvoyance; if it turns out to be false, she cites this as evidence of interference. No matter what happens, then, the clairvoyant's claim to be clairvoyant cannot be refuted. Her claim to clairvoyance is as empty and vacuous as the Perfect Sage's prediction.

Positions that are set up in this way so that nothing can possibly refute them are called *self-sealers*. A self-sealing position is one that is so constructed that no evidence can possibly be brought against it no matter what happens. This shows its vacuity, and it is precisely for this reason that we reject it.

People do not usually hold self-sealing positions in a blatant way; they tend to back into them. A person who holds that the American economy is controlled by an international Jewish conspiracy will point out people of Jewish extraction (or with Jewish names) who occupy important positions in financial institutions. This at least counts as evidence, though very weak evidence. And there seems to be much stronger evidence on the other side: There are a great many people in these institutions who are not Jews. To counter this claim, the person now argues that many of these other people are secretly Jews or are tools of the Jewish conspiracy. The Jews have allowed some non-Jews to hold important positions in order to conceal their conspiracy. What evidence is there for this? Well, none really, but that only helps prove how sneaky the Jewish conspiracy is. At this point, the position has become self-sealing, for all evidence cited against the existence of the conspiracy will be converted into evidence for its cleverness.

Ideologies and worldviews tend to be self-sealing. The Marxist ideology sometimes has this quality. If you fail to see the truth of the Marxist ideology, that just shows that your social consciousness has not been raised. The very fact that you reject the Marxist ideology shows that you are not yet capable of understanding it and that you are in need of reeducation. This is perfect self-sealing. Sometimes psychoanalytic theory gets involved in this same kind of self-sealing. People who vigorously disagree with certain psychoanalytic claims can be accused of repressing these facts. If a boy denies that he wants to murder his father and sleep with his mother, this itself can be taken as evidence of the strength of these desires and of his unwillingness to acknowledge them. If this kind of reasoning gets out of hand, then psychoanalytic theory also becomes self-sealing and empty. Freud was aware of this danger and warned against it.

So far, we have seen two ways in which an argument can be self-sealing: (1) It can invent an ad hoc or arbitrary way of dismissing every possible criticism. The clairvoyant can always point to interfering conditions without

going to the trouble of saying what they are. The anti-Semite can always cite Jewish cleverness to explain away counterevidence. We might call this self-sealing *by universal discounting*. (2) A theory can also counter criticism by attacking its critics. Critics of Marxism are charged with having a decadent bourgeois consciousness that blinds them to the facts of class conflict. The critic's response to psychoanalytic theory is analyzed (and then dismissed) as repression, a reaction formation, or something similar. Here self-sealing is achieved through an ad hominem fallacy. We might call this self-sealing *by going upstairs*, because the theorist is looking down on the critic.

Yet another form of self-sealing is this: (3) Words are used in such a way that a position becomes true *by definition*. For example, a person makes the strong claim that all human actions are selfish. This is an interesting remark, but it seems to be false, for it is easy to think of cases in which people have acted in self-sacrificing ways. To counter these obvious objections, the argument takes the following turn: When a person acts in a self-sacrificing way, what that person *wants* to do is help another even at her own expense. This is her desire or her motive, and that is what she acts to fulfill. So the action is selfish after all, because the person is acting to achieve what she wants. This is a self-sealing move, for it will not help to cite any behavior—even heroic self-destructive behavior—as counterevidence. If a person desires to do something even if it involves the sacrifice of her life, then she acts to fulfill her desire, and the act is again called *selfish*.

It is not hard to see the trick here. The arguer has chosen to use the word “selfish” in a new and peculiar way: A person is said to act selfishly if she does what she desires to do. This is not what we usually mean by this word. We ordinarily say that a person acts selfishly if she is too much concerned with her own interests at the expense of the interests of others. On this standard use of the word “selfish,” there are any number of counterexamples to the claim that all human actions are selfish. But these counterexamples do not apply when the word “selfish” is used in a new way, where “acting selfishly” comes close to meaning just “acting.” The point is that under this new meaning of “selfish,” it becomes empty (or almost empty) to say that all human actions are selfish. Thus, under one interpretation (the ordinary interpretation), the claim that all human actions are selfish is interesting but false. Under another interpretation (an extraordinary interpretation), the claim is true but vacuous. The position gets all its *apparent* interest and plausibility from a rapid two-step back-and-forth between these positions.

Self-sealing arguments often change their form under pressure in this way. A person will begin by holding a significant position that implies that facts are one way rather than another, but under the pressure of criticism will self-seal the position so that no evidence can possibly count against it. A theory that vacillates in this way is either *vacuous* or *false*. It is vacuous if self-sealing, false if not.

One good strategy for responding to this trick is to begin by charging a person who uses such an argument with saying something trivial, vacuous,

or boring. If, to meet this charge, he or she says something quite specific and important, then argument can proceed along normal lines. But it is not always easy to pin people down in this way. This becomes clear if you examine an argument between a Marxist and an anti-Marxist, between a psychoanalyst and a critic of psychoanalysis, or between individuals with different religious views. Their positions are often sealed against objections from each other, and then their arguments are almost always at cross-purposes.

Although we have emphasized how large-scale ideologies can become self-sealing, small-scale claims in everyday life are also often sealed against any possible refutation. In fact, a number of common words are used to this end. If someone says, "All true conservatives support school prayer," and a critic points out a conservative who opposes school prayer, then the original claim might be defended by saying, "He is not *truly* (or *really*) a conservative." If this response is trotted out in every case, it turns out that the original claim does not exclude anything. Similarly, the claim that "some students need to work harder than others, but if any student works hard enough, he or she will get good grades" can be protected simply by declaring that any student who works hard but does not get good grades does not work hard *enough*. Finally, someone who says, "If you think it over thoroughly, you will agree with me" can dismiss anyone who disagrees simply by denying that he thought it over *thoroughly*. Of course, these terms—"true," "real," "thorough(ly)," and "enough"—do not always make positions self-sealing. Nonetheless, these and other common terms are often used to seal positions against any possible criticism. When these terms are used in these ways, the resulting positions are empty and can be criticized in the same ways as self-sealing ideologies.

#### DISCUSSION QUESTIONS

1. Antony Flew famously wrote:

Now it often seems to people who are not religious as if there was no conceivable event or series of events the occurrence of which would be admitted by sophisticated religious people to be a sufficient reason for conceding "There wasn't a God after all" or "God does not really love us then." Someone tells us that God loves us as a father loves his children. We are reassured. But then we see a child dying of inoperable cancer of the throat. His earthly father is driven frantic in his efforts to help, but his Heavenly Father shows no obvious sign of concern. Some qualification is made—God's love is "not merely human love" or it is "an inscrutable love," perhaps—and we realize that such offerings are quite compatible with the truth of the assertion that "God loves us as a father (but, of course . . .)." We are reassured again. But then perhaps we ask: what is this assurance of God's (appropriately qualified) love worth, what is this apparent guarantee really a guarantee against? Just what would have to happen not merely (morally and wrongly) to tempt us but also (logically and rightly) to entitle us to say "God does not love us" or even "God does not exist"?<sup>2</sup>

(continued)

How would you answer Flew's question? If the answer to Flew's question were that nothing could entitle us to say this, as Flew suggests, then would this show that religious positions like this are self-sealing? That they are empty? Why or why not?

2. Some creationist critics of Darwin's theory of natural selection argue as follows:

Natural selection is a tautologous concept (circular reasoning) because it simply requires the fittest organisms to leave the most offspring and at the same time it identifies the fittest organisms as those that leave the most offspring. Thus natural selection seemingly does not provide a testable explanation of how mutation would produce more fit organisms.<sup>3</sup>

Does this argument show that Darwin's theory is self-sealing? How could defenders of natural selection best respond?

## NOTES

<sup>1</sup> John Stuart Mill, *A System of Logic* (London, 1843), book 2, chapter 3, section 2.

<sup>2</sup> Antony Flew, "Theology and Falsification," in *New Essays in Philosophical Theology*, ed. A. Flew and A. MacIntyre (New York: Macmillan, 1955), 98–99.

<sup>3</sup> Duane T. Gish, Richard B. Bliss, and Wendell R. Bird, "Summary of Scientific Evidence for Creation," *Impact*, May/June 1981, 95–96.

## REFUTATION

**C**hapter 1 showed how arguments can be used for justification and for explanation, but arguments can also be used for another purpose: refutation. This chapter will explain the nature of refutation and explore some of the main ways in which arguments can refute another argument or claim. These methods of refutation include counterexamples, *reductio ad absurdum*, and parallel reasoning. This last kind of refutation can reveal a large variety of fallacies in addition to those studied in previous chapters.

### WHAT IS REFUTATION?

---

In addition to justifying and explaining their conclusions, arguments are also sometimes used to refute other arguments. To *refute* an argument is to show that it is no good. Some writers, however, incorrectly use the term “refute” to mean something much weaker. They say such things as that Bill Clinton refuted the charges brought against him (by those attempting to impeach him while he was president), meaning nothing more than that he rejected or replied to the charges. This, however, is not what the word “refute” means. To refute the charges brought against him, Clinton would have to *give reason to believe* that these charges were erroneous. Refuting a charge requires giving an adequate argument against it. This takes a lot more work than simply denying it.

On the other hand, it is also important to remember that we can refute an argument without proving that its conclusion is false. A refutation of an argument is sufficient if it raises objections that cannot be answered. Consequently, the patterns of successful refutations mirror the criteria for a good argument, because the point of a refutation is to show that one of these criteria has not been met. Refutations, then, take four main forms: (1) We can argue that some of the premises are dubious or even false. (2) We can argue that the conclusion of the argument leads to absurd results. (3) We can show that the conclusion does not follow from the premises (or, in the case of an inductive argument, that the premises do not provide strong enough support for the conclusion). (4) We can show that the argument begs the question. This last charge was discussed in Chapter 16, so here we will focus on the first three methods of refutation.

**DISCUSSION QUESTION**

Is refuting an argument the same as justifying a belief that its conclusion is false? Is it the same as justifying a belief that the argument is invalid or weak? Why or why not?

**COUNTEREXAMPLES**

The first main way to attack an argument is to challenge one of its premises. We can argue that there is no good reason to accept a particular premise as true, asking, for example, “How do you know that?” If the premise is not justified, then the argument fails to justify its conclusion. More strongly, we can argue that the premise is actually false. In this second case, we refute an argument by refuting one of its premises.

One common way to refute a premise by showing that it is false is by producing a counterexample. Counterexamples are typically aimed at universal claims. This is true because a *single* contrary instance will show that a universal claim is false. If someone claims that *all* snakes lay eggs, then pointing out that rattlesnakes bear their young alive is sufficient to refute this universal claim. If the person retreats to the weaker claim that *most* snakes lay eggs, the guarding term makes it much harder to refute the claim. A single example of a snake that bears its young alive is not enough to refute this claim; we would have to show that a majority of snakes do not lay eggs. Here, instead of trying to refute the statement, we may ask the person to produce his *argument* on behalf of it. We can then attack this argument. Finally, if the person retreats to the very weak claim that at least *some* snakes lay eggs, then this statement becomes very difficult to refute. Even if it were false (which it is not), to show this we would have to check every single snake and establish that it does not lay eggs. So, as a rough-and-ready rule, we can say that the stronger a statement is, the more subject it is to refutation; the weaker it is, the less subject it is to refutation.

When a universal claim is refuted by a single case, that case is a *counterexample* to the universal claim. The pattern of reasoning is perfectly simple: To refute a claim that *everything* of a certain kind has a certain feature, we need find only *one* thing of that kind lacking that feature. In response to a counterexample, many people just repeat the misleading saying, “That’s the exception that proves the rule.” What most people do not realize is that “proves” originally meant “tests,” so all this saying means is that an apparent exception can be used to test a rule or a universal claim. When the exception is a true counterexample, the universal claim fails the test.

There are only two ways to defend a universal claim against a purported counterexample. Because the universal claim says that all things

of a certain kind have a certain feature, (1) one can deny that the apparent counterexample really is a thing of that kind, or (2) one can deny that the supposed counterexample really lacks that feature. For example, a defender of the claim that all snakes lay eggs might deny (1) that rattlesnakes are snakes, or (2) that rattlesnakes bear their young alive. Neither of these responses is plausible in this case. That is what makes this counterexample *decisive*.

Other counterexamples are not decisive. Indeed, some purported counterexamples miss their targets entirely. If a person claims that all snakes except rattlesnakes lay eggs, someone might respond with another counterexample: male snakes. This counterexample does not really refute the intended claim, since that claim was meant to be about the methods by which female snakes of various species give birth when they do give birth.

When a counterexample can be answered with a simple clarification or modification that does not affect the basic force of the original claim, it is a *shallow* counterexample. A *deep* counterexample is one that requires the original claim to be modified in more important or interesting ways. Shallow counterexamples can sometimes be fun as jokes, but they are usually not of much help in refuting arguments, since basically the same argument can be resurrected in a slightly different form. Indeed, people who give too many shallow counterexamples can be annoying. If you really want to understand a subject matter, you should look for counterexamples that are deep.

Ethics is an area where arguments often turn on counterexamples. Consider the traditional moral precept “Do unto others as you would have them do unto you.” This principle captures an important moral insight, but, if taken quite literally, it is also subject to counterexamples. After a defendant is found guilty of murder, a judge sentences the defendant to prison, even though the judge would not want to be treated that way herself—she would not want to be sentenced to prison. The Golden Rule was not intended to rule out the judge’s behavior, though it seems to. Conversely, a sadomasochist enjoys beating other people. When asked whether he would like to be treated in that way, he replies, “Yes.” It is obvious that the Golden Rule was not intended to approve of the sadomasochist’s behavior, but it seems to. The task, then, is to reformulate the Golden Rule so as to avoid these (and similar) counterexamples. That is not as easy as it might seem.

No discussion of counterexamples is complete without a mention of the Morgenbesser retort. Though the exact story is now shrouded in the mists of time, it has come down to us from the 1950s in the following form: In a lecture, a British philosopher remarked that he knew of many languages in which a double negative means an affirmative, but not one language in which a double affirmative means a negative. From the back of the room came Sydney Morgenbesser’s retort: “Yeah, yeah.”

**EXERCISE I**

Find a counterexample to each of the following claims, if possible.

Example: *Claim:* "Sugar" is the only word in which an s is pronounced *sh*.

*Counterexample:* Oh, sure.

1. No prime number is even.
2. Three points always determine a plane.
3. Balloons that are filled with helium always rise in the air.
4. All mammals bear their young live.
5. You can never get too much of a good thing.
6. What you don't know can't hurt you.
7. You can't be too careful.
8. It is always wrong to tell a lie.
9. You should never ask someone else to do something that you are not willing to do yourself.
10. If lots of people do something, then it must not be wrong for me to do it.
11. If it would be horrible for everyone to do something, then it would be morally wrong for anyone to do it.
12. If it would not be horrible for everyone to do something, then it would not be morally wrong for anyone to do it.

**EXERCISE II**

There cannot possibly be any counterexamples to the following claims. Explain why.

1. There is life on the moon.
2. Killing is usually wrong.
3. Any short person is a person.
4. Every horse is an animal.
5.  $2 + 2 = 4$ .
6. Everything with a size also has a shape.

**DISCUSSION QUESTIONS**

1. How can the Golden Rule be best reformulated to avoid the above counterexamples of the judge and the sadomasochist? Can you think of any counterexamples to this reformulation of the Golden Rule?

2. Is the Morgenbesser retort a shallow counterexample or a deep counterexample? Why?
3. When theologians claim that God can do anything, atheists sometimes respond that God cannot make a stone that is so large that God cannot lift it, or that God cannot make a circle with four sides. Are these really counterexamples to the theologians' claim? Why or why not?

## REDUCTIO AD ABSURDUM

Particular counterexamples can normally be used to refute claims only if those claims are universal, so how can we refute claims that are not universal? One method is to show that the claim to be refuted implies something that is ridiculous or absurd in ways that are independent of any particular counterexample. This mode of refutation is called a *reductio ad absurdum*, which means a reduction to absurdity. Reductios, as they are called for short, can refute many different kinds of propositions. They are sometimes directed at a premise in an argument, but they can also be used to refute a conclusion. This method of refutation will not show exactly *what* is wrong with the argument for that conclusion, but it will show that *something* is wrong with the argument, because it cannot be sound if its conclusion is false. That might be enough in some situations.

For example, suppose someone argues that because there is a tallest mountain and a heaviest human, there must also be a largest integer. We might respond by arguing as follows: Suppose there is a largest integer. Call it  $N$ . Since  $N$  is an integer,  $N + 1$  is also an integer. Moreover,  $N + 1$  is larger than  $N$ . But it is absurd to think that any integer is larger than the largest integer. Therefore, our supposition—that there is a largest integer—must be false.

In this mathematical example a contradiction is derived, but absurdity also comes in other forms. Suppose a neighbor tells a parent, "The local public schools are so bad that you ought to send your kids to private school," and the parent responds, "Do you think I'm rich?" The point of this rhetorical question is that it is absurd to think that the parent is rich, presumably because of her lifestyle or house, which the neighbor can easily see. Without being rich, the parent cannot afford a private school, so the neighbor's advice is useless.

Often the absurdity is derived indirectly. A wonderful example occurred in the English parliamentary debate on capital punishment. One member of Parliament was defending the death penalty on the grounds that the alternative—life in prison—was much more cruel than death. This claim was met with the following reply: On this view, those found guilty of first-degree murder ought to be given life in prison, and the death penalty should

be given to those who commit some lesser offense. The first speaker could respond in several ways, because this *reductio* depends on background assumptions that the first speaker could question. First, he might deny that the most severe crime should receive the most severe penalty possible. If the first speaker sees life in prison as too cruel to be inflicted on anyone, then he might call for the abolition of life imprisonment and keep the death penalty as the most severe punishment. Alternatively, the first speaker could claim that, even though the death penalty is less severe than life in prison, it is still fitting in some other way for the most severe crime, first-degree murder. Finally, of course, the first speaker could simply *accept* the supposedly absurd result and apply life imprisonment to first-degree murder, while using the death penalty for lesser crimes. In fact, however, the first speaker was unwilling to accept any of these alternatives. He simply tried a rhetorical trick and got caught.

These *reductios* are fairly good, but other *reductios* fail for a variety of reasons. To succeed in refuting a claim, a *reductio ad absurdum* argument must meet two main requirements. First, the result must really be *absurd*. Often opponents try to reduce a view to absurdity but really only draw out implications of the view that are not absurd at all. For example, in a famous debate in which Thomas Huxley defended a theory of evolution, Bishop Wilberforce asked Huxley whether he had descended from apes on his mother's side or on his father's side of the family. This question was intended to draw laughter from the crowd, and it did, partly because they and Wilberforce thought that any answer to the question would be absurd. Nonetheless, Huxley could respond that he had descended from apes on both sides of his family. Because that response was not really absurd—regardless of how absurd it seemed to Wilberforce—the bishop's attempt did not really refute Huxley's claim.

In other cases, one cannot deny that a certain result really would be absurd, but the *reductio* still fails because the claim to be refuted does not really *imply* that absurdity. For example, opponents sometimes say that the theory of evolution implies that animals are constantly evolving, so they cannot be divided into separate species. This would be absurd, because it is easy to observe distinct species. The theory of evolution, however, does not really imply this absurdity, so this *reductio* fails to refute that theory. It fails to meet the second requirement for successful *reductios*, which is that the claim to be refuted must actually imply the absurdity.

Finally, it is important to notice that *reductios* can be deep or shallow in much the same way as counterexamples. Sometimes a claim really does imply a result that is absurd, but it can be modified in some minor way so as to avoid the absurd result. For example, if a fan says, "Tiger Woods is better than any golfer ever," someone might respond that Woods is himself a golfer, so this claim implies that Woods is better than himself, which is absurd. Of course, the fan meant to say, "Woods is better than any *other* golfer ever," so this *reductio* is shallow. The *reductio* does refute the original form of the claim, but the main force of the claim is restored by the minor modification.

A *reductio ad absurdum* is deep only if it reveals that a claim implies an absurd result that cannot be avoided without modifying the claim in essential respects or giving it up entirely.

In sum, then, a *reductio ad absurdum* argument tries to show that one claim, *X*, is false because it implies another claim, *Y*, that is absurd. To evaluate such an argument, the following questions should be asked:

1. Is *Y* really absurd?
2. Does *X* really imply *Y*?
3. Can *X* be modified in some minor way so that it no longer implies *Y*?

If either of the first two questions is answered in the negative, then the *reductio* fails; if the third question receives an affirmative answer, then the *reductio* is shallow. Otherwise, the *reductio ad absurdum* argument is both successful and deep.

### EXERCISE III

Evaluate the following *reductio ad absurdum* arguments by asking the above three questions.

1. **CLAIM TO BE REFUTED:** Even the worst of enemies can become friends.  
**REDUCTIO:** If people are enemies, then they are not friends. If they do become friends, then they are not enemies. So it's absurd to think that enemies can be friends.
2. **CLAIM TO BE REFUTED:** This ball is both red all over and green all over.  
**REDUCTIO:** If it is red, it reflects light within a certain range of wavelengths. If it is green, it reflects light within a different range of wavelengths. These ranges do not overlap, so it is absurd to think that anything can reflect both kinds of light. Thus, a ball cannot be both red and green all over.
3. **CLAIM TO BE REFUTED:** Most children in Lake Wobegon are above average (in intelligence).  
**REDUCTIO:** If so, the average (intelligence) would really be higher than it is; and then it would not be true that most children in Lake Wobegon are above the real average (intelligence).
4. **ARGUMENT TO BE REFUTED:** Your brain is mostly empty space, because the subatomic particles in it are very far apart.  
**REDUCTIO:** That's absurd, because my brain is solid, and it works pretty well.
5. **CLAIM TO BE REFUTED:** Some things are inconceivable.  
**REDUCTIO:** Consider something that is inconceivable. Since you are considering it, you are conceiving it. But then it is conceivable as well as inconceivable. That is absurd. So nothing is inconceivable.

**EXERCISE IV**

Spell out a reductio ad absurdum argument to refute each of the following claims. If no such reductio is possible, explain why.

1. Some sisters are nephews.
2. Some fathers were never children.
3. Most students scored better than the median grade on the last test.
4. Almost everyone in this class is exceptional.
5. There is an exception to every universal claim.
6. I know that I do not know anything.
7. Some morally wrong actions are morally permitted.
8. God exists outside of time, and we will meet Him someday.
9. There is a male barber in this town who shaves all and only the men in this town who do not shave themselves. (Hint: Does he shave himself?)
10. Most of the sentences in this exercise are true.

**DISCUSSION QUESTIONS**

1. The legal case of *Plessy v. Ferguson*, 163 U.S. 537 (1896), questioned the constitutionality of a law requiring racial segregation in railroad cars. Opponents of the law gave the following reductio argument. How could defenders of segregation respond to this argument? Is their response adequate? Is any response adequate? Why or why not?

The same argument that will justify the state legislature in requiring railways to provide separate accommodations for the two races will also authorize them to require separate cars to be provided for people whose hair is of a certain color, or who are aliens, or who belong to certain nationalities, or to enact laws requiring colored people to walk upon one side of the street, and white people upon the other, or requiring white men's houses to be painted white, and colored men's black, or their vehicles or business signs to be of different colors, upon the theory that one side of the street is as good as the other, or that a house or vehicle of one color is as good as one of another color.

2. Many atheists try to refute belief in God with the following reductio ad absurdum argument: God is defined to be all-good and all-powerful (as well as all-knowing). If God is all-good, then God prevents as much evil as He can. If God is all-powerful (and all-knowing), then God can prevent all evil. Thus, if a traditional God did exist, there would be no evil in the world. But that's absurd. There is obviously lots of evil in the world. Therefore, God does not exist. Evaluate this reductio argument. How could religious believers best respond?

## STRAW MEN AND FALSE DICHOTOMIES

Very often when trying to refute either by counterexample or by reductio, people move too quickly. The general rule is this: Before trying to refute someone's claim, it is important to make sure that you understand his or her position. If you misunderstand what your opponent is claiming, but you go ahead and attack a specific claim anyway, then the claim you attack will not be the claim that your opponent made. You might even fail to refute any position that anyone ever really held. This is called the fallacy of *attacking a straw man*.

Sometimes people attack a straw man intentionally as part of a rhetorical strategy. They mischaracterize their opponents' position on purpose in order to make their opponents look silly by associating their opponents with a claim that really is silly. For example, if someone proposes providing free condoms in high schools as a way to reduce teen pregnancy and sexually transmitted diseases, an opponent might respond, "I suppose you want high school students to have more sex!" Of course, that is not what was proposed. Similarly, if someone supports loosening restrictions on marijuana, an opponent might reply, "Do you really want everyone to get high all the time? Do you want our society to be filled with drug addicts?" This attack on a straw man might fool some listeners, perhaps by fueling their fears, but it does not refute any real arguments for legalizing marijuana.

The fallacy of attacking a straw man can also arise from an honest mistake. Some people get so wrapped up in their own arguments that they forget the view against which they are arguing. The opponent can also be partly to blame. If someone states her position obscurely, it might not be clear whether the speaker would go so far as to make a certain claim. Then someone might attack that further claim, honestly believing that the speaker had adopted it. Alternatively, a critic might refute that further claim simply to make the speaker clarify her position by explicitly saying that, that is not what she meant to say. In such ways, it might be useful to refute a position that the speaker does not really hold, even though, of course, doing so does not refute any position that the speaker actually does hold.

In more insidious cases, straw men are often set up by means of a related fallacy—*false dichotomy*. With regard to the Iraq war, President Bush often said something like this: "I had a choice to make: Either take the word of a madman [Saddam Hussein] or defend America. Given that choice, I will defend America every time." The crucial phrase, of course, is "given that choice." If those were the only options, then Bush's critics would also defend America every time. The problem lies in Bush's suggestion that his opponents do not want to defend America and would instead "take the word of a madman." That insinuation sets up a straw man.

Political rhetoric is filled with such false dichotomies that set up straw men. A cable news host is reported to have said, "Sure, it's not great having

the government collect our telephone records, but it is better than having them collect our body parts." Let's hope there is a third option! Opponents of the government collecting phone records, of course, think that there is another way to avoid having to collect body parts. But then they sometimes add, "Either you are opposed to the government collecting phone records or you don't care about civil rights." This is just a false dichotomy on the other side. Those who favor the government collecting phone records do care about civil rights, even though they favor some intrusions (which they see as minor intrusions) on those rights in order to fight terrorism.

False dichotomies like these are parodied by Stephen Colbert when he says, "Either you are with us or you are with the terrorists," "Either you're for the war [in Iraq] or you hate America," and his best, "George W. Bush: great president or the greatest president?" The trick here is obviously to give you some choice, so that you end up committed to the option that you choose, but your choices are limited to ones that Colbert gives you.

After listing his dichotomy, Colbert usually adds, "It's that simple!" The best response is to recognize that these issues are usually not simple at all. These tricks work partly because many people long for simple choices or they fail to notice any third option. Whenever someone tells you that you have only two alternatives, you should look carefully for other possibilities. And whenever your choice among the options seems obvious, you should ask whether the rejected options have been set up as straw men rather than characterized fairly.

#### EXERCISE V

Do the following arguments attack straw men? Why or why not?

1. Anyone who thinks that the United States should not have sent troops to Iraq must think that the suffering Saddam Hussein inflicted on his own citizens doesn't matter much.
2. Anyone who thinks that the United States should have sent troops to Iraq must think that they will only be there for a short time.
3. Humans could not have been created in the image of God, because God is not a physical being, and only physical beings can have images.
4. Atheists think that God does not exist, so everything is permitted. But even atheists must admit that I would not be permitted to kill them! So atheism is nonsense.
5. The theory of evolution says that humans are no different from apes, but humans are clearly smarter than apes, so the theory of evolution must be wrong.
6. Stephen Colbert again: "Evolutionists' main claim is that one day we decided to stop being monkeys and turned ourselves into humans. Well, if that's true, why aren't more monkeys escaping from zoos? Think about it. They could turn into humans, then disguise themselves as janitors and walk out of their cages. But I guess evolution doesn't have an answer for that one."<sup>1</sup>

**DISCUSSION QUESTION**

Find five more examples of attacking straw men in your local newspaper, in a talk show on television, or in a college course.

**REFUTATION BY PARALLEL REASONING**

Even if its premises and conclusion cannot be refuted by any counterexample or reductio, a deductive argument can also be refuted by showing that it is invalid. We know that an argument is not valid if it starts from true premises and leads to a false conclusion. Often, however, we cannot point this out to refute an argument, because the truth or falsity of the conclusion is the very thing at issue. When this problem arises, a typical device is to point out that by arguing in the same way, or a similar way, we can reach a conclusion that is unsatisfactory.

Here is a simple example:

**CARY:** Most of the people in this class are college students. Most college students study hard. Therefore, most of the people in this class study hard.

**DAVID:** That's just like arguing that most whales live in the sea, and most animals that live in the sea are fish, so most whales are fish.

At first sight, it might not be clear how the second argument could show anything about the first argument. What do whales have to do with students? The point, however, is simply that the two arguments share a basic form: Most As are Bs, and most Bs are Cs, so most As are Cs. Thus, if the second argument is not valid by virtue of that form, then the first is also not valid by virtue of that same form. The second argument is obviously not valid, since its premises are true but its conclusion is false. This shows that the first argument is not valid, at least by virtue of this shared form. Even though the first argument still might be valid on some other basis, its defenders at least owe an alternative account of its validity. Often there will be none.

Refuting an argument by showing that it is *just like* another argument that is obviously no good is a common device in everyday discussions. Here's another example:

**MATTHEW:** If I had a higher salary, I could buy more things; so, if everyone had higher salaries, everyone could buy more things.

**KIRSTY:** That's just like arguing that, if one person stands up at a ball game, he will get a better view; so, if everyone stands up, everyone will get a better view.

At first sight, it may not be obvious whether Matthew's style of reasoning is valid or not. Kirsty's response shows that Matthew's argument is invalid by providing an instance in which the same style of reasoning takes us

from something true to something that is obviously false, because, if everyone stands up at a ball game, only the tallest people will be able to see better. Kirsty's response also shows *why* Matthew's argument is invalid: Just as one person's ability to see can be affected by other people standing up, because this raises the height that is necessary to see, so one person's ability to buy can be affected by other people having more money, if this raises prices and thereby raises the amount of money that is necessary to buy things.

This fallacy is often called a *fallacy of composition*, because it rests on the mistaken assumption that what is true of the parts is also true of the whole that is composed out of those parts. Each person in a class has a mother, but the whole class does not have a mother. The earth might be heating up on average in the long run even if some locations on earth have a cool summer one year. These obvious mistakes can be cited to show that and why potentially misleading arguments with the same form are no better. Lots of new fallacies can be revealed in this way by deploying this method of refutation by parallel reasoning.

Of course, not every refutation of this kind is so simple or so successful. To understand the criteria that must be met for such a refutation to work, it will be useful to consider a more complex example that reveals some of the ways to respond to a charge of "That's just like arguing. . . ." The example concerns proposed legal restrictions on gun ownership. The National Rifle Association (NRA) feared that these restrictions would lead to a total ban on guns, which they opposed, so they widely distributed a bumper sticker that read:

- (1) If guns are outlawed, only outlaws will have guns.

The point, presumably, was that most people would add the suppressed premise:

- (2) It would be bad if only outlaws had guns,

and then reach the conclusion:

- (3) Therefore, guns should not be outlawed.

This argument is not completely clear, partly because it is not clear who counts as an "outlaw." Some critics poke fun at this bumper sticker because (1) seems true by definition if outlaws include anyone who breaks any law, because anyone with a gun breaks a law if guns are outlawed. But what the NRA probably means by "outlaws" are people who commit violent crimes, such as robbery and murder. It is not strictly true that these will be the only people with guns if guns are outlawed, since police and some present gun owners would keep their guns. Nonetheless, these exceptions do not touch the NRA's main claim, which is that law-abiding people who would give up their guns if guns were outlawed would then not have guns to defend themselves against violent criminals.

How can an opponent try to refute this argument? There are several possibilities, but what defenders of gun control in fact did was distribute other bumper stickers. One of them read:

(1\*) If gum is outlawed, only outlaws will have gum.

The main point might be just to parody the NRA bumper sticker, but, if we take it more seriously, (1\*) also suggests an application of the method of refutation by parallel reasoning. The parallel argument would continue like this:

(2\*) It would be bad if only outlaws had gum.

(3\*) Therefore, gum should not be outlawed.

This conclusion, however, is not obviously false. Indeed, it seems true: People should be allowed to chew gum. Moreover, (2\*) seems false, because nothing particularly bad would happen if only outlaws chewed gum. For these reasons, this bumper sticker cannot really refute the original argument. This failure illustrates two general tests: A refutation by parallel reasoning works only if the conclusion of the parallel argument really is unacceptable and only if the premises of the parallel argument really are true.

But opponents of the NRA did not stop there. They distributed a third bumper sticker:

(1\*\*) If guns are outlawed, only outlaws will shoot their children by mistake.

The argument behind this new bumper sticker is again not clear. If it is a straightforward instance of refutation by parallel reasoning, then the parallel argument would add, "It would be bad if only outlaws shot their children by mistake" and conclude, "Guns should not be outlawed." But that is the very conclusion the NRA wants to reach; so this newest argument could not refute the original one. Nonetheless, a different argument might lie behind this third bumper sticker. The point seems to be that gun owners sometimes shoot their children by mistake, and we can minimize such tragedies by reducing the number of gun owners through laws against guns. The argument then runs something like this:

(1\*\*) If guns are outlawed, only outlaws will shoot their children by mistake.

(2\*\*) It would be good if outlaws were the only ones who shot their children by mistake.

∴(3\*\*) Therefore, guns should be outlawed.

This conclusion would seem false to the NRA, and this argument might also seem to suggest that the same form of reasoning could lead to opposite conclusions: (3) and (3\*\*). Moreover, (2\*\*) seems true. This premise does not say that it is good for outlaws to shoot their children by mistake. Instead, it says

that it would be good if nobody else shot their children by mistake. So far, so good. Notice, however, that this latest argument, (1\*\*)–(3\*\*), does not have the same form as the original argument, (1)–(3), because (2\*\*) and (3\*\*) are about what is good and what ought to be law, whereas (2) and (3) are about what is bad and what ought *not* to be law. (3) includes a negation that is missing in (3\*\*).

The next question is whether this disanalogy is *important*. If not, the latest bumper sticker still might refute the original one. The NRA, however, might argue that this difference *is* important. The fact that a law has bad effects overall *does* show that the law should not be passed, whereas the fact that a law would have good effects overall is *not* enough to show that the law should be passed, since the law still might violate individual rights that cannot be overridden by good effects on others. This claim is controversial, but, if it can be defended, then this parallel argument, (1\*\*)–(3\*\*), fails to show that the original argument, (1)–(3), is invalid. More generally, then, a refutation by parallel reasoning works only if the two arguments really do have relevantly similar structures—that is, only if one argument really is *just like* the other in relevant respects.

In sum, the method of refutation by parallel reasoning can be used to show that an argument is invalid by presenting another argument with essentially the same form in which the inference takes us from obvious truths to an obvious falsehood. In response to such an attack, a defender of the original argument has three main options. The defender might

1. deny that the conclusion of the parallel argument is false,
2. deny that the premises of the parallel argument are true, or
3. deny that the supposedly parallel argument really has essentially the same form as the original argument.

If any of these responses is justified, then the attempt to refute the original argument by parallel reasoning fails.

This procedure is admittedly imprecise. There will sometimes be disputes about whether the premises of the parallel argument really are true, or clearly true, and whether the conclusion of the parallel argument really is false, or clearly false. Moreover, we have given no general explanation of the notion that two arguments have the *same basic form*. Some forms of argument were discussed in previous chapters, but they are only part of the story. We have not discussed and cannot discuss all possible forms of argument. Yet it remains a fact that people can often see that two arguments have the same essential form and, through seeing this, decide that an argument presented to them is invalid. This ability is the basis of sound logical judgment. It is also the basis of wit. It is at best mildly funny to say that if God had wanted us to fly, He would have given us wings. You have to be fairly clever to reply at once, “If God had wanted us to stay on the ground, He would have given us roots.”

## EXERCISE VI

In each of the following examples, does the parallel argument succeed in refuting the original argument? Why or why not? Consider the three possible responses listed above. If the original argument is refuted, is there some simple way to fix it so that it cannot be refuted by this parallel reasoning? If so, how? (You might try to add a premise whose analogue would be false in the parallel argument.)

1. **CHRIS:** The United States is wealthy, so its citizens are wealthy as well.  
**PAT:** That's just like arguing that a building is expensive, so the nails in its walls are expensive as well.
2. **CHAUVINIST:** Since women are the only people who can bear children, they should bear children.  
**FEMINIST:** That's like arguing that, if I am the only person who can wiggle my own ears, then I should wiggle my ears.
3. **NEWT:** Orphanages are fine places, as the movie *Boys Town* shows.  
**CRITIC:** That's just like saying that Oz is a fine place, as *The Wizard of Oz* shows.
4. **A YOUNG CHILD AT 6:00 AM:** It's morning. Morning is the time to wake up. So it's time to wake up.  
**THE CHILD'S SLEEPY PARENT:** That's just like arguing that it's daytime, and daytime is the time to eat lunch, so it's time to eat lunch.
5. **MARK:** You shouldn't walk on that grass, because if everybody did that, the grass would die.  
**BOB:** That's just like arguing that I shouldn't go to this movie right now, because if everybody did that, the theater would be packed like a can of sardines.
6. **THOMAS:** Everything in the world has a cause, so the world itself must have a cause.  
**TONY:** That's just like arguing that every leg in the relay race was run by a single runner, so the entire race itself must have been run by a single runner.
7. **HAWK:** Nuclear deterrence must work, because we have never had a nuclear exchange as long as we have maintained nuclear deterrence.  
**DOVE:** That's just like arguing that hanging garlic by the front door must keep thieves away, because I put garlic there and my house has never been robbed.
8. **LIBERAL:** We ought to provide condoms for high school students, because they are going to have sex anyway.  
**CONSERVATIVE:** That's just like arguing that we should provide high school students with guns, because they are going to use guns anyway.
9. **SCIENTIST:** My initial steps toward human cloning, no matter how controversial, are important because they bring the debate out into the public.

(continued)

**OPPONENT:** That's just like arguing that we should start a fire in a house in order to bring the debate about getting a new fire engine out into the public. (paraphrased from an actual interview on National Public Radio)

10. **KING:** "In your statement you asserted that our actions, even though peaceful, must be condemned because they precipitate violence. But can this assertion be logically made? Isn't this like condemning the robbed man because his possession of money precipitated the evil act of robbery?" (from Martin Luther King's "Letter from Birmingham Jail")
11. **A:** He owns a red car, so he owns a car.  
**B:** That's just like arguing that he owns a toy duck, so he owns a duck.
12. **A:** He is holding a baby girl, so he is holding a baby.  
**B:** That's just like arguing that he is driving a fire truck, so he is driving a fire.

#### EXERCISE VII

For each of the following arguments, find another argument with the same basic form in which the premise or premises are clearly true and the conclusion is clearly false.

1. If tea is dangerous, so is coffee. Tea isn't dangerous. So coffee isn't either.
2. If it were about to rain, it would be cloudy. It is cloudy. So it's about to rain.
3. Fred had either ice cream or cake for dessert. He had cake. So he must not have had ice cream.
4. You cannot pass laws against dangerous drugs, because there is no way to draw a sharp line between dangerous and nondangerous drugs.
5. If you have never written a novel, then you are in no position to make judgments about novels. So don't presume to criticize mine.
6. Since I have written several novels, I am in a position to know which novels are good. So you ought to trust me when I say that this one is great.
7. There's nothing wrong with smoking, since the longer you smoke, the longer you live.
8. If one has nothing to hide, one should not be afraid of being investigated. So no one should object to being investigated.
9. Radicals should not be granted freedom of speech, because they deny this freedom to others.
10. In nature, a species is more likely to survive when its weak members die out, so we should let the weak in our society die out.

11. Buses use more gas than cars, so the city cannot reduce gas consumption by providing more buses.
12. Boxing can't be very bad, since so many people like it.
13. This war is just, for to say otherwise in public would be to aid our enemies.
14. If you don't buy the most expensive shoes, you buy cheap ones. You don't want cheap shoes. So you should buy the most expensive shoes.
15. I'd rather be smart than strong, so I am going to quit exercising and spend all day in the library.
16. You don't want to be this murderer's next victim, so you had better convict her and send her to prison where she can't hurt you.
17. If it weren't for America, these refugees would have nowhere to go; so they should adopt the American way of life and give up their old culture.
18. You can't be right, because, if the answer were that obvious, someone would have thought of it before.

### NOTE

<sup>1</sup>Stephen Colbert, *I Am America (And So Can You!)* (New York: Grand Central Publishing, 2007), 198.



## AREAS OF ARGUMENTATION

*Particular subject matters often give rise to new kinds of arguments and special standards for assessing arguments. One reason is that arguments in different areas have different purposes. Much scientific reasoning is aimed at explaining observations, including findings in experiments, whereas the point of moral reasoning is not to explain what happens but, rather, to evaluate and to determine which acts are justified. The audience also varies from one area of argumentation to another. Some religious reasoning is addressed only to people who already accept certain basic claims of a religious tradition, such as that the Bible is inspired, whereas philosophical reasoning typically prides itself in questioning common assumptions and authorities. Similarly, arguments about different topics often take place within different institutional contexts with distinctive rules. Legal arguments, for example, often assume specified burdens of proof, certain basic laws (such as in a constitution), and other fundamental features of the legal system in the jurisdiction, whereas moral reasoning is sometimes used to overturn those very same features of legal systems. Because these areas of argumentation differ in so many ways, we need to look at each of them separately and carefully to see which standards are at play in each area. That is one main goal of Part V. Another goal is to provide extensive examples of arguments both as targets for criticism and as models of how to construct a good argument. We can get better at arguing for our own views if we learn to appreciate the strengths in arguments given by other people, including our opponents. Toward these ends, Part V will discuss and give illustrations of legal, moral, scientific, religious, and philosophical reasoning in Chapters 18–22, respectively.*



## LEGAL REASONING

*Law is a prime example of how institutional setting affects standards of argumentation, for legal systems have their own procedures and authorities. Legal decisions are often based on statutes, constitutions, and precedents that are peculiar to a particular legal system. These sources of law are often vague, and they can conflict, so they need to be interpreted. This chapter will discuss legal interpretation and legal reasoning in general. The general forms of legal arguments will then be illustrated by looking in more detail at the law of discrimination and affirmative action.*

Legal decisions have concrete effects on people's lives. In criminal cases, judges can deprive people of their freedom or even their lives. In civil cases, judges often take away large sums of money, the custody of children, and so on. Constitutional decisions can affect the basic rights of all citizens, even those who have never been in a courtroom. These decisions are made because certain legal arguments are accepted and others are rejected.

Unfortunately, it is sometimes difficult to find any good reason for a legal decision. Crucial facts might not be known, and the law is sometimes unclear or inconsistent. Some cases "fall between the cracks," so no law seems to apply. Human beings have a remarkable ability to produce weird cases that would tax the wisdom of Solomon.

Even in the toughest cases, judges and juries must reach *some* decision. Outside the law, we can often just let matters ride—we can postpone a decision until further facts are established, or even declare that the issues are too vague to admit of any decision. This is rarely an option in a legal case. If *A* sues *B*, either *A* or *B* must win. The judge cannot say, "This case is too tough for me. I'm not going to rule on it." Throwing the case out of court amounts to ruling in favor of the defendant. A decision must be made, usually in a relatively short period of time.

These pressures have led lawyers and judges to develop many ingenious ways to argue. Lawyers cite statutes, precedents, and historical contexts. They claim the authority of common sense and science, and they cite scholarly articles, even some by philosophers. They deploy metaphors and rhetorical devices—almost anything to convince the judge or jury to decide in favor of their clients. The variety of these arguments makes legal reasoning complex and also fascinating.

Despite this variety, some rough generalizations can be made: A decision in a legal case usually depends on (1) questions of fact and (2) questions of law.

## COMPONENTS OF LEGAL REASONING

### QUESTIONS OF FACT

A *criminal law* prohibits a certain kind of behavior and assigns a punishment to those who violate it. When a person is accused of violating this law, a trial is held to determine whether *in fact* he or she has done so. The judge instructs the jury on the law bearing on the case. If the members of the jury then decide that the accused has violated the law, they find him or her guilty, and the judge usually hands down the punishment stated in the statute.

In a *civil* suit, one party sues another, say, for breach of contract. Because states have laws governing contracts, once more a trial is held to decide whether *in fact* there was a contract (instead of some other speech act, as in *Hawkins v. McGee*, cited on p. 28) and whether *in fact* it was breached. If a breach of contract is found, the judge or jury awards damages as the law specifies.

Although questions of fact arise in all cases, criminal and civil cases do differ in the *burden of proof*—in who is required to establish the facts and to what degree of certainty. In a criminal procedure, the prosecution must establish its case *beyond a reasonable doubt*. In a civil case, the burden of proof is less. Generally, the case is won by the party who shows that the *preponderance* of evidence favors his or her side of the case. Although this is a bit too simple, it is sometimes said that if the scales tip ever so slightly in favor of *A* rather than *B*, then *A* wins the case.

The only way to carry the burden of proof is to present *evidence*. This evidence can contain conflicts and unclarities, which make it hard to prove the facts. Sometimes the facts are so complex that they simply cannot be proved one way or the other, and sometimes the distinction between facts and law is not so clear.

These problems arise often in cases that raise larger social issues. For example, in the famous case of *Brown v. Board of Education* (347 U.S. 483 [1954]), which found segregated schools unconstitutional, the Supreme Court answered the question of law at issue by saying “the opportunity of an education . . . is a right which must be available to all on equal terms.” The Court next asked: “Does segregation of children in public schools solely on the basis of race . . . deprive the children of the minority group of equal educational opportunities?” This question was presented as a question of fact. The Court answered in the affirmative and tried to justify its answer by citing various psychological studies of the performance of minority children from segregated schools. This answer would be accepted by most people today, but the studies used as proof were controversial and inconclusive, so the Court had to decide whether studies of this kind were reliable enough to serve as evidence

in this case. Moreover, the answer to the above question also depends on what counts as “equal educational opportunities” for the purposes of the law. For example, the studies cited by the Court found that segregated schools “affect the motivation of a child to learn,” but these factual studies could not determine whether lowered *motivation* to learn counts as lowered *opportunity* to learn. The Court had to decide this issue because it in effect determines what the law is—what it prohibits and what it allows. Thus, what was presented as a question of fact turned out to be at least partly a question of law. In such cases, it is not clear where law ends and facts begin.

## QUESTIONS OF LAW

Even after the facts are determined, no decision can be reached without determining what the law is. The law varies from place to place and from time to time, so we have to know what the law is at the right time and place. This is determined mainly by looking at the legal institutions that actually exist. In our legal system, there are three main sources of the law: statutes, the Constitution, and precedents.

**STATUTES.** Roughly, statutes are general rules of law passed by legislatures. Statutes are made at various levels (federal, state, and local), and they cover various subjects, including crimes as well as property, contracts, and other areas of civil law. There are also statutes governing the procedures and kinds of evidence that can be presented in court.

When a general statute is applied to a particular case, the legal argument is often primarily deductive. For example, Sally drove ninety-five miles per hour in front of Hanover High School at 4 PM on a school day. It is illegal to drive over fifteen miles per hour in front of any school at 4 PM on a school day. Therefore, Sally’s driving was illegal. Of course, there are lots of suppressed premises, such as that ninety-five miles per hour is over fifteen miles per hour. Other assumptions are trickier: Sally might not be found guilty if she had an excuse or justification, such as that a terrorist held a gun to her head. Still, it is often obvious that Sally had no excuse or justification. If we add this claim as a premise, then, with a little fiddling, the legal argument against Sally can be made deductively sound.

Such simple cases are common, but they are also boring. Things get much more difficult and interesting when a statute is *vague*, so that it is not clear whether the statute applies to the case at issue. Then the statute must be *interpreted*. We need some way to tell more precisely what the law prohibits and what it allows.

The first step in interpreting a statute is to look carefully at the *words* in the statute and their literal *meanings*. But the courts must often look beyond the mere words of the statute. This need arises when the words are unclear and when they lead to absurd results. For example, suppose a city council passes an ordinance requiring zoos to provide clean, dry cages for all mammals.

This works fine until one zoo puts a whale in its aquarium. The whale would be in trouble if the courts stuck to the words of the ordinance. Fortunately, the courts can also consider the *intentions* of the legislators, which can be gleaned from their debates about the law. Of course, the city council might not have thought at all about whales, or they might have thought that whales are fish instead of mammals. Thus, if their intentions are what the legislators consciously had in mind, then we also need to consider the deeper, more general *purpose* of the legislators—the goal they were trying to reach or the moral outlook they were trying to express. This purpose is revealed by the wider historical context and by other laws made by the same legislature. In our example, the purpose of the statute was obviously to provide a healthy environment for mammals in zoos. This purpose is best served by interpreting the ordinance so that it does not require dry cages for whales.

In addition to words, intentions, and purposes, *moral beliefs* are often used to interpret statutes. Judges often argue that a statute should be interpreted one way by claiming that any other interpretation would lead to some kind of practical difficulty or moral unfairness. Such arguments are effective when everyone agrees about what is immoral or unfair, but judges often depend on more controversial moral beliefs. Critics claim that judges should not use their own moral views in this way, but there is no doubt that many or even all judges do in fact reason from such moral premises.

It should be clear that none of these methods of interpretation is mechanical, and none guarantees a single best interpretation of every statute. Part of the legal controversy is often over which factors can or should be used to argue for an interpretation. When all is said and done, legal reasoning from statutes is often far from the straightforward deduction that it appears to be in simple cases.

**THE CONSTITUTION.** Even when a statute has been interpreted, it is sometimes not clear whether the statute is *valid*—whether it has any legal force. This is determined by the Constitution. The Constitution occupies a special place in the legal system of the United States. If any statute conflicts with the Constitution, including its amendments, that statute has no legal force. Generally it is not the role of courts to enact laws, but the courts do have the power to strike down laws if they conflict with provisions in our Constitution.

It is easy to imagine clear cases of laws that violate constitutional provisions. If the state of Kansas began printing its own money, that would plainly violate the constitutional provision that reserves this right to the federal government. But, typically, those constitutional questions that reach the courts are not clear-cut. Even more so than statutes, provisions in the Constitution are very general and sometimes vague. This vagueness serves a purpose: The framers of the Constitution recognized that they could not foresee every eventuality, so they wanted to allow future courts to interpret the Constitution as cases arose. But the vagueness of the Constitution also creates problems. Interpretations can often conflict and fuel controversy. As with statutes, arguments for and against interpretations of the Constitution usually refer

to the words of the Constitution, the intentions and purposes of the framers of the Constitution, the effects of adopting an interpretation, moral beliefs, and so on. Such arguments are often inconclusive. The Supreme Court is then the final arbiter on questions of constitutionality.

*PRECEDENTS.* The U. S. legal system is not only a constitutional system; it is also partly a system of common law. This means that lawyers and judges often cite precedents in arguments for present decisions. A *precedent* is simply a past case or decision that is supposed to be similar to the present case.

The practice of citing precedents might seem strange at first sight. Why should one case provide any reason for a decision in a different case? The answer is that the cases resemble each other in important respects. Of course, when there is an important enough difference between the cases, they should be *distinguished*, and then the precedent provides no argument in the present case. But, when there is no important enough difference, like cases should be treated alike. If similar precedents were not followed, the legal system would lack continuity, and this would make it unfair and ineffective. Of course, past decisions that were mistaken or immoral should not be continued. That is why precedents can be *overturned*. Nonetheless, our legal system assumes that, if there is no adequate reason to overturn a precedent or to distinguish the precedent from the present case, then the precedent provides some reason to decide the present case in the same way as the precedent. This general doctrine of precedent is often called *stare decisis*—to adhere to previous decisions.

Precedents are used for many different purposes. When a statute is vague, precedents are often used to argue for one interpretation over another. When no statute applies directly, precedents are often used to argue about what the law is. Precedents can also be used in arguments for general questions of fact, or simply as sources of persuasive rhetoric.

The form of arguments from precedents also varies. Often a judge or lawyer merely quotes part of the opinion in the precedent and treats that quotation as an authoritative pronouncement of the law. Arguments from precedents are then similar to arguments from legislative statutes, and there often arises a similar need to interpret the judicial pronouncement in the precedent. In other precedents, the judge chooses to make the decision without explicitly formulating any general rule of law. The precedent can still be used to argue for future decisions by emphasizing analogies and discounting differences between the precedent and the present case.

One example occurs in the case of *Plessy v. Ferguson*, 163 U.S. 537 (1896). Louisiana passed a statute that required blacks and whites to use “separate but equal” cars in trains. Plessy refused to comply, and he claimed that the Louisiana law violated the Fourteenth Amendment to the Constitution, which forbids states to deprive anyone of “the equal protection of the laws.”

In his argument for this claim, Plessy cited the precedent of *Yick Wo v. Hopkins*, 118 U.S. 356 (1886). That case was about an ordinance in San Francisco that

required a permit from the Board of Supervisors for any public laundry not operated in a brick or stone building. On its face, this ordinance was supposed simply to prevent fires. In practice, however, the Board of Supervisors granted permits to all but one of the non-Chinese applicants and denied permits to all the Chinese applicants. Because of this practice, *Yick Wo* claimed that the ordinance violated the equal protection clause. The Supreme Court agreed and declared the ordinance unconstitutional, at least insofar as it gave the city power to grant and refuse permits “without regard to the competency of the persons applying, or the propriety of the places selected for the carrying on of business.”

The argument from a precedent to a decision in a present case is often presented as an argument from *analogy*. In this form, the argument emphasizes similarities between the cases, and then concludes that the decision in the present case should be the same as in the precedent. *Plessy*’s argument then appears to run something like this:

- (1) The ordinance in *Yick Wo* was declared unconstitutional.
  - (2) The ordinance in *Yick Wo* is similar to the statute in *Plessy* in several respects.
- ∴ (3) The statute in *Plessy* also ought to be declared unconstitutional.

This argument is not very good as it stands, so we need to add some suppressed premises.

The first step is to construct a list of the respects in which the cases are similar. That is not always easy. When we are evaluating someone else’s argument, we can focus on the similarities that he or she mentions. But when we are constructing our own legal arguments, we have to be more creative; we have to formulate the respects in which the cases are supposed to be similar. Some similarities do not matter: It is clearly irrelevant that the laws in *Yick Wo* and *Plessy* both contain more than ten words or that both apply to large cities. This much is assumed by both sides in the case, and legal reasoning would be impossible without assuming that many such similarities are irrelevant. Likewise, not all differences matter: It is not important, even if true, that *Yick Wo* was married and over fifty years old, but *Plessy* was not. To discount or distinguish the precedent, one must show that some difference between *Yick Wo* and *Plessy* is important enough to justify reaching different decisions in these cases. The central question, then, asks which factors—similarities and differences—do matter. In general, the answer is that a factor is relevant when it is needed to justify the decision in the precedent. These factors are often called the *ratio decidendi*—the reason for the decision.

Incorporating the doctrine of precedents as a suppressed premise, the argument from *Yick Wo* to *Plessy* can now be reconstructed as follows:

- (1) The ordinance in *Yick Wo* was declared unconstitutional.
- (2) The ordinance in *Yick Wo* is similar to the statute in *Plessy* in several respects (*A, B, C, D*, and so on).

- (3) These are the features that justified declaring the ordinance in *Yick Wo* unconstitutional.
- (4) There are no important enough differences between *Yick Wo* and *Plessy* to justify distinguishing the precedent.
- (5) *Yick Wo* ought not to be overturned.
- (6) If a precedent is similar to a present case in the respects that justified the decision in the precedent, and if the precedent ought not to be either overturned or distinguished, then the present case ought to be decided in the same way as the precedent.  
∴(7) The statute in *Plessy* ought to be declared unconstitutional.

This argument is now valid, but validity does not get us very far. We still need to know whether its premises are true.

Clearly, the crucial question is this: How do we determine which features of the precedent are needed to justify that decision? What we need to do is to extract a general rule of law that provides the best justification for the precedent decision.

The most obvious way to argue that a certain feature is important is to look at the written opinion in the precedent and see what the Court said—more specifically, what reasons it gave for its decision. In *Yick Wo*, the Court wrote:

whatever may have been the intent of the ordinances as adopted, . . . though the law itself be fair on its face and impartial in appearance, yet, if it is applied and administered by public authority with an evil eye and an unequal hand, so as practically to make unjust and illegal discriminations between persons in similar circumstances, material to their rights, the denial of equal justice is still within the prohibition of the Constitution.

Here the Court explicitly announces that the intent of the ordinance and its appearance (for example, whether the ordinance explicitly mentions race or ethnic background) did *not* matter to their decision. The Court also declares that it *did* matter that the ordinance in practice creates inequalities in rights. Such official pronouncements by a court have considerable force for future courts in legal arguments.

Another way to determine which factors matter is to apply the necessary condition test or the sufficient condition test from Chapter 10 to a group of precedents. These tests had to be passed by each side in the *Plessy* case. *Plessy* claimed that a sufficient condition of unconstitutionality is that a law has a discriminatory *effect* on the rights of a particular racial or ethnic group. This claim would fail the sufficient condition test if there were any precedent still in force in which a law was found to have a discriminatory effect but the law was not held unconstitutional. Because there was no such precedent, the sufficient condition test does not exclude *Plessy*'s claim that discriminatory effect is sufficient by itself to make a law unconstitutional.

On the other side, the Court claimed that discriminatory effect is not sufficient, because discriminatory *motive* is a necessary condition for a law to be unconstitutional under the equal protection clause. For this claim to pass the necessary condition test, there must have been no precedent still in force in which a law was held unconstitutional under the equal protection clause but the Court did not find any discriminatory motive. *Plessy* claimed that *Yick Wo* was such a case, but the Court responded that, even if those who passed the ordinance had no discriminatory motive, the administration of the ordinance in *Yick Wo* “was held to be a covert attempt on the part of the municipality to make an arbitrary and unjust discrimination against the Chinese race.” If so, the necessary condition test does not rule out the Court’s claim that discriminatory motive is necessary for unconstitutionality in this case.

This disagreement reveals the limits on the tests of necessary conditions and sufficient conditions. These tests are useful when there is a rich body of coherent precedents. But when there are not enough precedents of the right kinds, and when the precedents are not coherent, the necessary condition test and the sufficient condition test cannot be used to rule out conflicting interpretations of the precedents.

When the actual precedents are not enough, judges sometimes refer to *hypothetical cases*. In *Plessy*, a judge might imagine a law with discriminatory effect but no discriminatory motive. If the judge can show why this law should be found unconstitutional, this would suggest that a discriminatory motive is not really necessary for a violation of the equal protection clause. This takes some imagination, and it also requires judges to apply their moral beliefs about which laws should be allowed. Some critics deny that moral arguments should have any legal force, because they are so controversial. Nonetheless, there is no doubt that judges often do in fact assume such moral beliefs in arguments from precedents.

A final point to remember is that arguments from precedents are usually defeasible, like other inductive arguments (see Chapter 8). One reason is that more precedents might be found, and these new precedents might conflict with the precedents in the original argument. Another reason is that any precedent can be overturned. Precedents are not supposed to be overturned unless they are very badly mistaken or immoral, but this is always a possibility. Nonetheless, even though arguments from precedents always might be refuted in such ways, precedents can still provide reasons for legal decisions.

So far we have looked at arguments from precedents as ways to determine what is necessary or sufficient to violate the law. Even after this is determined, the law still must be applied to the facts in the present case. In *Plessy*, the Court held that there was no intent to discriminate, because the statute in *Plessy* was “reasonable” and “enacted in good faith for the promotion of the public good and not for the annoyance or oppression of a particular class.” This claim is highly questionable. In his famous dissent, Justice Harlan denies it when he writes:

Everyone knows that the statute in question had its origin in the purpose, not so much to exclude white persons from railroad cars occupied by blacks, as to exclude colored people from coaches occupied by or assigned to white persons. . . . No one would be so wanting in candor as to assert the contrary.

If Harlan is right, the Court's argument has a false premise, so the statute in *Plessy* should have been found unconstitutional even if the Court was right about what was necessary to find a law unconstitutional.

We can summarize this discussion by listing various ways in which arguments from precedents can fail:

1. The precedent and the present case might not *truly* resemble each other in the ways that the argument claims.
2. The respects in which the cases resemble each other might not be *important* enough to justify the same decision in the present case.
3. The precedent and the present case might also differ from each other in important respects that justify *distinguishing* the precedent.
4. The precedent might be mistaken or immoral enough to be *overturned*.
5. There might be other, stronger precedents that *conflict* with the precedent in the argument.

Whenever you evaluate or present any argument from a precedent, you need to ask whether the argument fails in any of these ways.

## THE LAW OF DISCRIMINATION

These general methods of legal reasoning can be seen at work in a particular area of constitutional law—the law of discrimination. To understand the cases in this area, some background will be helpful.

### THE EQUAL PROTECTION CLAUSE

The provision of the Constitution that governs discrimination is the equal protection clause of the Fourteenth Amendment. It provides as follows:

No state shall make or enforce any law which shall . . . deny to any person within its jurisdiction the equal protection of the laws.

The clearest thing about this clause is that it is not clear. Whatever it means, it cannot mean that laws cannot ever treat people unequally. Criminal laws treat those who commit crimes quite differently from those who do not. The general idea behind the clause seems to be that like cases should be treated in like ways. Put negatively, the clause prohibits unequal treatment when there is no significant difference. This, however, is still both general and vague, for we need principles that determine what sorts of likenesses matter and what kinds of differences are significant.

Going back to the historical context in which the Fourteenth Amendment was adopted, we know that it was *intended* to prohibit unequal treatment on the basis of “race, color, or previous condition of servitude” (a phrase that occurs in the companion Fifteenth Amendment on voting rights). More specifically, it was one of those constitutional provisions intended to protect the newly emancipated slaves. This was the primary *purpose* of these provisions, but the *language* is more general, giving like protection to all citizens of the United States.

### APPLYING THE EQUAL PROTECTION CLAUSE

After the Fourteenth Amendment was adopted, many questions arose concerning its interpretation and application. The amendment explicitly refers only to state laws, but the state does many things besides pass laws, so the courts had to determine what counts as a *state action*. In a series of cases, the amendment was interpreted to mean that only positive actions of the state fell under the amendment. Thus, when thugs broke up a black political rally, with the police standing by doing nothing to protect the demonstrators, the Supreme Court ruled that this was not a violation of the equal protection clause because the state itself had not participated in the action (*U.S. v. Cruikshank*, 1875). On this view, the state was forbidden to aid discrimination, but it was not required to protect anyone against it.

Another issue that arose concerned what the state has to do to justify treating people differently. Here the courts decided that it was not their business to examine the details of legislation to make sure that the laws were as equitable as possible. The task of making laws, they held, falls to legislatures, and the courts gave legislatures wide latitude in formulating these laws. Flagrant violations of the equal protection clause could lead to the decision that the law was unconstitutional, but only if the law failed what became known as the *rational-relation test*. This test required only that the unequal treatment of individuals be reasonably likely to achieve some legitimate end.

A final issue was whether the Fourteenth Amendment protected only the civil rights of citizens or also rights of other kinds. In *Strauder v. West Virginia* (1880), a law that made blacks ineligible for jury duty was struck down on the grounds that the equal protection clause prohibits discrimination in areas of civil rights. In *Yick Wo v. Hopkins* (1886), the Court applied the equal protection clause to discrimination in areas of economic rights. *Yick Wo* also established that the equal protection clause protects not only blacks but also other groups (such as Chinese) and that laws that do not explicitly mention racial or ethnic groups can violate the equal protection clause if they are applied unequally in practice.

In 1896, the Supreme Court decided the case of *Plessy v. Ferguson*, which was about a Louisiana statute enforcing racial segregation in public transportation. This was clearly a state action, and the Court continued to apply the rational-relation test. The main issues were whether the equal protection clause extends to social rights, whether the segregation law served any

reasonable purpose, and whether the separate facilities were truly equal. The Court held that the segregation statute did not violate the equal protection clause so long as the facilities were equal. This became known as the *separate-but-equal doctrine*.

### THE STRICT SCRUTINY TEST

As soon as *Plessy* was decided, southern and some border states rapidly passed a whole series of segregation laws. These were subsequently upheld by the courts on the precedent of *Plessy* and its separate-but-equal doctrine. The result was the introduction of a system of racial segregation throughout much of the country.

The doctrine of the rational-relation test remained basically unchanged until the 1940s. During World War II, the Supreme Court had to decide whether it was constitutional to relocate Japanese Americans away from Pacific ports. In *Korematsu v. United States*, 323 U.S. 214 at 216 (1944), the Court announced that

all legal restrictions which curtail the rights of a single racial group are immediately suspect. That is not to say that all such restrictions are unconstitutional. It is to say that the courts must subject them to the most rigid scrutiny.

It is ironic that the Court did not strike down the Japanese relocation orders, but these cases established a new interpretation of equal protection that eventually greatly increased the power of the courts to strike down discriminatory laws.

On this new interpretation of the equal protection clause, most laws still need to pass only the rational-relation test, but there are two features of a law that serve as *triggers* of strict scrutiny. A law must pass *strict scrutiny* if the law either restricts a *fundamental right* or employs a *suspect classification*. Fundamental rights concern such things as the right to vote or the right to procreate. A classification is suspect if it concerns race, religion, national origin, and so on. Under the new interpretation of the equal protection clause, states could still pass laws restricting fundamental rights, and these laws could still employ suspect classifications, but, when they did so, a heavy burden of proof fell on them. To justify such a law, the state had to show that (1) the legislation serves a *legitimate* and *compelling* state interest, and that (2) it does so in the *least intrusive* way possible (or, as later courts put it, the legislation must be “narrowly tailored” to fit the interest). This is the *strict-scrutiny test*.

It should be clear that the rational-relation test is easy to meet, whereas the test of strict scrutiny is nearly impossible to satisfy. It is not hard to show that a piece of legislation has some chance of serving some legitimate goal and, thus, passes the rational relation test. It is difficult to show that the purpose of a law is compelling—that is, of overwhelming importance; and it is even more difficult to show that the stated purpose cannot be achieved by any less intrusive means. Thus, in adopting this new test, the Court no longer showed great deference to state legislatures, as it did in *Plessy*, when it applied the rational-relation test. Instead, the heaviest burden was shifted

to the states in areas that involved what the Court declared to be suspect classifications or fundamental rights.

This new test was crucial for the decision of *Brown v. Board of Education* (1954), which declared segregation in public schools unconstitutional. The *Brown* opinion does not directly mention strict scrutiny, but this test looms in the background. Segregation clearly involves a suspect classification, but the Court emphasizes, “the opportunity of an education . . . is a *right* which must be available to all on equal terms” (emphasis added). The next step is to argue that segregated schools violate this right by their very nature, even if all “tangible” factors are equal. This violation of a fundamental right triggers strict scrutiny, and the *Brown* opinion then simply assumes that segregation in education will fail this test. Separate but equal is thus found unconstitutional, at least in education, and *Plessy* is in effect overturned.

### THE BAKKE CASE

After *Brown*, the Supreme Court struck down segregation in many other areas—transportation, parks, libraries, and so on—as well as laws against racial intermarriage (though not until 1967, after Barack Obama’s parents were married). Another string of decisions required states to use busing as a means to end segregation in school systems. The Court also required some employers to hire or promote minimum percentages of minorities to overcome the effects of illegal discrimination in employment.

In response to these court decisions, some schools and companies voluntarily took steps to overcome what they saw as the effects of past discrimination. These steps required them to use racial classifications, and that raised the issue of reverse discrimination. Part of the issue was about what to call such programs. Their opponents label them “reverse discrimination,” but their defenders refer to them as “affirmative action.” Both names clearly involve evaluation. A more neutral description might be “preferential treatment.”

The classic case in this area is *Regents of the University of California v. Bakke* (1978) (hereafter, *Bakke*). The basic situation was that the medical school of the University of California at Davis had very few minority students, so they created a special admissions program that set aside sixteen seats for minorities who were disadvantaged. Bakke, who was not a member of any of the specified minorities, applied to the school but was rejected even though he had higher scores on admissions tests than some minority members who were admitted under the special admissions program.

Bakke claimed that Davis’s special admissions program violated the equal protection clause of the U.S. Constitution, the California Constitution, and Title VI of the Civil Rights Act of 1964, which provides:

No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.

The Supreme Court was thus asked to rule on four main issues:

1. Did the Davis special admissions program violate the equal protection clause of the Constitution?
2. Does reference to race without judicial findings of particular past discrimination violate this constitutional guarantee?
3. Did the Davis special admissions program violate Title VI of the Civil Rights Act?
4. Should Davis be required to admit Bakke into its medical school?

The decision of the Supreme Court on these issues is so complicated that it takes a scorecard to follow it:

Question Ruled On	Brennan, Marshall, Blackmun, and White	Powell	Burger, Stewart, Stevens, and Rehnquist	The Majority
1	No	Yes	No decision	No decision
2	No	No	No decision	No
3	No	Yes	Yes	Yes
4	No	Yes	Yes	Yes

Because the justices split into two groups of four, the remaining justice, Powell, determined the majority on most issues. Powell, however, was the only justice who argued that the Davis program was unconstitutional. Four others (Brennan, Marshall, Blackmun, and White) dissented. The remaining four (Burger, Stevens, Stewart, and Rehnquist) chose not to address this constitutional issue because they had already ruled out the Davis program under Title VI. Since a majority did not join Powell in his opinion, the Court did *not* explicitly declare the Davis program unconstitutional. But the Davis program was held to violate Title VI, and Davis was ordered to admit Bakke, because Burger, Stevens, Stewart, and Rehnquist did join Powell on these issues. Despite Davis's loss, the Court took the opposite position on the second issue. Powell argued that it was not always unconstitutional for the state to refer to race, and he created a majority when he was joined by Brennan, Marshall, Blackmun, and White. Thus, each group of justices got part of what it wanted.

The constitutional issues are raised most directly in the opinions of Powell and Brennan (excerpted below). These opinions differ not only in their conclusions but also in their interpretations of the equal protection clause. Powell argued that the Davis program and any consideration of race must be subjected to the test of strict scrutiny. He held that the Davis program did not meet the high standards of strict scrutiny, but some other consideration of race might, such as in the Harvard College admissions program, which used race as a "goal" rather than as a "quota" (much like the University of Michigan Law School admissions program in the *Grutter* case below).

In contrast, Brennan argued for a new interpretation of the equal protection clause. On this new interpretation, strict scrutiny would still be applied to most racial classifications, but strict scrutiny need not be passed when the state uses a racial classification to serve a *benign, remedial purpose*. The purpose of a racial classification is benign when it does not stigmatize anyone and was not adopted out of any discriminatory motive, and it is remedial if the state used the racial classification because the state found that, without the racial classification, an underprivileged group would suffer harm or differential impact because of past discrimination in society at large.

Powell criticized Brennan's conditions on the grounds that the notion of stigma is too vague and that it is not groups but individuals who are protected by the equal protection clause. Brennan responded by distinguishing stigma from other harms and by emphasizing the importance of groups. The heart of the controversy was, thus, about the conditions under which to apply strict scrutiny.

**REGENTS OF THE UNIVERSITY  
OF CALIFORNIA V. BAKKE**

(438 U.S. 268, 1978)

*Excerpts from Justice Powell's opinion:*

Racial and ethnic classifications . . . are subject to stringent examination without regard to . . . additional characteristics. We declared as much in the first cases explicitly to recognize racial distinctions as suspect: ". . . [A]ll legal restrictions which curtail the rights of a single racial group are immediately suspect. That is not to say that all such restrictions are unconstitutional. It is to say that courts must subject them to the most rigid scrutiny" (*Korematsu*, 323 U.S. 214 at 216 [1944]). The Court has never questioned the validity of those pronouncements. . . .

Petitioner urges us to adopt for the first time a more restrictive view of the Equal Protection Clause and hold that discrimination against members of the white "majority" cannot be suspect if its purpose can be characterized as "benign." The clock of our liberties, however, cannot be turned back to 1868. It is far too late to argue that the guarantee of equal protection to *all* persons permits the recognition of special wards entitled to a degree of protection greater than that accorded others.

Moreover, there are serious problems of justice connected with the idea of preference itself. First, it may not always be clear that a so-called preference is in fact benign. Courts may be asked to validate burdens imposed upon individual members of particular groups in order to advance the group's general interest. . . . Nothing in the Constitution supports the notion that

individuals may be asked to suffer otherwise impermissible burdens in order to enhance the societal standing of their ethnic groups. Second, preferential programs may only reinforce common stereotypes holding that certain groups are unable to achieve success without special protection based on a factor having no relationship to individual worth. . . . Third, there is a measure of inequity in forcing innocent persons in respondent's position to bear the burdens of redressing grievances not of their making.

Petitioner contends that on several occasions this Court has approved preferential classifications without applying the most exacting scrutiny. Most of the cases upon which petitioner relies are drawn from three areas: school desegregation, employment discrimination, and sex discrimination. Each of the cases cited presented a situation materially different from the facts of this case. . . . [W]e have never approved preferential classifications in the absence of proven constitutional or statutory violations. . . . When a classification denies an individual opportunities or benefits enjoyed by others solely because of his race or ethnic background, it must be regarded as suspect. . . .

*Excerpts from Justice Brennan's opinion:*

. . . Unquestionably we have held that a government practice or statute which restricts "fundamental rights" or which contains "suspect classifications" is to be subjected to "strict scrutiny" and can be justified only if it furthers a compelling government purpose and, even then, only if no less restrictive alternative is available. . . . But no fundamental right is involved here. . . . Nor do whites as a class have any of the "traditional indicia of suspectness: the class is not saddled with such disabilities, or subjected to such a history of purposeful unequal treatment, or relegated to such a position of political powerlessness as to command extraordinary protection from the majoritarian political process." . . .

[The] fact that this case does not fit neatly into our prior analytic framework for race cases does not mean that it should be analyzed by applying the very loose rational-basis standard of review that is the very least that is always applied in equal protection cases. . . .

[B]ecause of the significant risk that racial classifications established for ostensibly benign purposes can be misused, causing effects not unlike those created by invidious classifications, it is inappropriate to inquire only whether there is any conceivable basis that might sustain such a classification. Instead, to justify such a classification an important and articulated purpose for its use must be shown. In addition, any statute must be stricken that stigmatizes any group or that singles out those least well represented in the political process to bear the brunt of a benign program. Thus our review under the Fourteenth Amendment should be strict—not "'strict' in theory and fatal in fact," because it is stigma that causes fatality—but strict and searching nonetheless. . . .

**DISCUSSION QUESTION**

Do you agree with Powell or with Brennan in their debate about whether strict scrutiny should be applied to affirmative actions programs in admissions? Why?

**LEGAL DEVELOPMENTS SINCE BAKKE**

The legal story of affirmative action gets very complicated after the *Bakke* decision. Because of the split within the Court, it was not clear which parts, if any, of Powell's opinion had force as precedent. It was also not clear which kinds of programs were close enough to the Harvard program that Powell endorsed; and it was not clear how to extend Powell's emphasis on diversity in education to other areas, such as employment and government contracts. As a result, some affirmative action programs were struck down and others were upheld while commentators scrambled to explain the underlying pattern and rationale for the Court's decisions.

The year after *Bakke*, in *United Steelworkers of America v. Weber*, 445 U.S. 193 (1979), the Supreme Court upheld a preferential treatment quota in employment. Later, in *Richmond v. J. A. Croson Co.*, 488 U.S. 469 (1989), the Court struck down a local program that set aside a specified percentage of municipal contracts for minority-controlled firms. Then, in *Metro Broadcasting v. FCC*, 497 U.S. 547 (1990), the Court allowed a Federal Communications Commission "distress sale" policy that permitted a broadcaster whose license or renewal application has been designated for a revocation hearing to sell the license to a buyer if, but only if, that buyer is at least 50 percent minority-owned. Of course, the majority opinions gave reasons for each decision, but dissents were common and fierce, and it was hard to see how the pieces of the puzzle fit together.

The law was somewhat clearer in college admissions. As Justice O'Connor put it, "Since this Court's splintered decision in *Bakke*, Justice Powell's opinion announcing the judgment of the Court has served as the touchstone for constitutional analysis of race-conscious admissions policies. Public and private universities across the nation have modelled their own admissions programs on Justice Powell's views on permissible race-conscious policies." Still, there was great uncertainty about exactly which kinds of admissions programs were allowed by Powell's reasoning in *Bakke*. Some critics also held that the precedent of *Bakke* should be overturned or disregarded, and the Court should hold, instead, that all race-conscious programs are unconstitutional.

This issue came to a head in 2003, when the Court considered two very different kinds of admissions programs practiced at the University of Michigan under its president, Lee Bollinger (hence the names of the following cases). Michigan's Law School admissions program was very much like the

Harvard program that Powell endorsed, so that program was upheld. The admissions program at Michigan's undergraduate College of Literature, Science, and the Arts, in contrast, gave a specific number of points to all applicants in underrepresented minority groups. Critics claimed that this program amounted to a "quota" more like the Davis Medical School admissions program that Powell had struck down in *Bakke*, and the Court struck it down. The differences between these programs thus define the line between what is and is not permitted in college admissions today.

## GRUTTER V. BOLLINGER

---

(539 U.S. 306, 2003)

*Syllabus:*

The University of Michigan Law School (Law School), one of the Nation's top law schools, follows an official admissions policy that seeks to achieve student body diversity through compliance with *Regents of Univ. of Cal. v. Bakke*. Focusing on students' academic ability coupled with a flexible assessment of their talents, experiences, and potential, the policy requires admissions officials to evaluate each applicant based on all the information available in the file, including a personal statement, letters of recommendation, an essay describing how the applicant will contribute to Law School life and diversity, and the applicant's undergraduate grade point average (GPA) and Law School Admissions Test (LSAT) score. Additionally, officials must look beyond grades and scores to so-called "soft variables," such as recommenders' enthusiasm, the quality of the undergraduate institution and the applicant's essay, and the areas and difficulty of undergraduate course selection. The policy does not define diversity solely in terms of racial and ethnic status and does not restrict the types of diversity contributions eligible for "substantial weight," but it does reaffirm the Law School's commitment to diversity with special reference to the inclusion of African-American, Hispanic, and Native-American students, who otherwise might not be represented in the student body in meaningful numbers. By enrolling a "critical mass" of underrepresented minority students, the policy seeks to ensure their ability to contribute to the Law School's character and to the legal profession.

When the Law School denied admission to petitioner Grutter, a white Michigan resident with a 3.8 GPA and 161 LSAT score, she filed this suit, alleging that respondents had discriminated against her on the basis of race in violation of the Fourteenth Amendment, Title VI of the Civil Rights Act of 1964, and 42 U.S.C. § 1981; that she was rejected because the Law School uses race as a "predominant" factor, giving applicants belonging to certain minority groups a significantly greater chance of admission than students with similar credentials from disfavored racial groups; and that respondents had no compelling interest to justify that use of race. . . .

*Justice O'Connor delivered the opinion of the Court, in which Justices Stevens, Souter, Ginsburg, and Breyer joined:*

... We have held that all racial classifications imposed by government "must be analyzed by a reviewing court under strict scrutiny." This means that such classifications are constitutional only if they are narrowly tailored to further compelling governmental interests. . . .

With these principles in mind, we turn to the question whether the Law School's use of race is justified by a compelling state interest. Before this Court, as they have throughout this litigation, respondents assert only one justification for their use of race in the admissions process: obtaining "the educational benefits that flow from a diverse student body." . . .

The Law School's educational judgment that such diversity is essential to its educational mission is one to which we defer. The Law School's assessment that diversity will, in fact, yield educational benefits is substantiated by respondents and their amici. Our scrutiny of the interest asserted by the Law School is no less strict for taking into account complex educational judgments in an area that lies primarily within the expertise of the university. Our holding today is in keeping with our tradition of giving a degree of deference to a university's academic decisions, within constitutionally prescribed limits. . . .

As part of its goal of "assembling a class that is both exceptionally academically qualified and broadly diverse," the Law School seeks to "enroll a 'critical mass' of minority students." The Law School's interest is not simply "to assure within its student body some specified percentage of a particular group merely because of its race or ethnic origin." That would amount to outright racial balancing, which is patently unconstitutional. Rather, the Law School's concept of critical mass is defined by reference to the educational benefits that diversity is designed to produce.

These benefits are substantial. As the District Court emphasized, the Law School's admissions policy promotes "cross-racial understanding," helps to break down racial stereotypes, and "enables [students] to better understand persons of different races." These benefits are "important and laudable," because "classroom discussion is livelier, more spirited, and simply more enlightening and interesting" when the students have "the greatest possible variety of backgrounds." . . .

Moreover, universities, and in particular, law schools, represent the training ground for a large number of our Nation's leaders. . . . In order to cultivate a set of leaders with legitimacy in the eyes of the citizenry, it is necessary that the path to leadership be visibly open to talented and qualified individuals of every race and ethnicity. All members of our heterogeneous society must have confidence in the openness and integrity of the educational institutions that provide this training. . . .

Even in the limited circumstance when drawing racial distinctions is permissible to further a compelling state interest, government is still

"constrained in how it may pursue that end: [T]he means chosen to accomplish the [government's] asserted purpose must be specifically and narrowly framed to accomplish that purpose." The purpose of the narrow tailoring requirement is to ensure that "the means chosen 'fit' . . . the compelling goal so closely that there is little or no possibility that the motive for the classification was illegitimate racial prejudice or stereotype." . . .

To be narrowly tailored, a race-conscious admissions program cannot use a quota system. . . . We are satisfied that the Law School's admissions program, like the Harvard plan described by Justice Powell, does not operate as a quota. Properly understood, a "quota" is a program in which a certain fixed number or proportion of opportunities are "reserved exclusively for certain minority groups." . . . The Law School's goal of attaining a critical mass of underrepresented minority students does not transform its program into a quota. As the Harvard plan described by Justice Powell recognized, there is of course "some relationship between numbers and achieving the benefits to be derived from a diverse student body, and between numbers and providing a reasonable environment for those students admitted." . . . Nor, as Justice Kennedy posits, does the Law School's consultation of the "daily reports," which keep track of the racial and ethnic composition of the class (as well as of residency and gender), "suggest[ ] there was no further attempt at individual review save for race itself" during the final stages of the admissions process. To the contrary, the Law School's admissions officers testified without contradiction that they never gave race any more or less weight based on the information contained in these reports. Moreover, . . . the number of African-American, Latino, and Native-American students in each class at the Law School varied from 13.5 to 20.1 percent, a range inconsistent with a quota.

That a race-conscious admissions program does not operate as a quota does not, by itself, satisfy the requirement of individualized consideration. When using race as a "plus" factor in university admissions, a university's admissions program must remain flexible enough to ensure that each applicant is evaluated as an individual and not in a way that makes an applicant's race or ethnicity the defining feature of his or her application. The importance of this individualized consideration in the context of a race-conscious admissions program is paramount. Here, the Law School engages in a highly individualized, holistic review of each applicant's file, giving serious consideration to all the ways an applicant might contribute to a diverse educational environment. The Law School affords this individualized consideration to applicants of all races. There is no policy, either *de jure* or *de facto*, of automatic acceptance or rejection based on any single "soft" variable. Unlike the program at issue in *Gratz v. Bollinger* [see below], the Law School awards no mechanical, predetermined diversity "bonuses" based on race or ethnicity. . . .

What is more, the Law School actually gives substantial weight to diversity factors besides race. The Law School frequently accepts nonminority applicants with grades and test scores lower than underrepresented

minority applicants (and other nonminority applicants) who are rejected. This shows that the Law School seriously weighs many other diversity factors besides race that can make a real and dispositive difference for nonminority applicants as well. By this flexible approach, the Law School sufficiently takes into account, in practice as well as in theory, a wide variety of characteristics besides race and ethnicity that contribute to a diverse student body. . . .

Petitioner and the United States argue that the Law School's plan is not narrowly tailored because race-neutral means exist to obtain the educational benefits of student body diversity that the Law School seeks. We disagree. Narrow tailoring does not require exhaustion of every conceivable race-neutral alternative. Nor does it require a university to choose between maintaining a reputation for excellence or fulfilling a commitment to provide educational opportunities to members of all racial groups. Narrow tailoring does, however, require serious, good faith consideration of workable race-neutral alternatives that will achieve the diversity the university seeks.

We agree with the Court of Appeals that the Law School sufficiently considered workable race-neutral alternatives. The District Court took the Law School to task for failing to consider race-neutral alternatives such as "using a lottery system" or "decreasing the emphasis for all applicants on undergraduate GPA and LSAT scores." But these alternatives would require a dramatic sacrifice of diversity, the academic quality of all admitted students, or both. . . . The United States advocates "percentage plans," recently adopted by public undergraduate institutions in Texas, Florida, and California to guarantee admission to all students above a certain class-rank threshold in every high school in the State. The United States does not, however, explain how such plans could work for graduate and professional schools. Moreover, even assuming such plans are race-neutral, they may preclude the university from conducting the individualized assessments necessary to assemble a student body that is not just racially diverse, but diverse along all the qualities valued by the university. We are satisfied that the Law School adequately considered race-neutral alternatives currently capable of producing a critical mass without forcing the Law School to abandon the academic selectivity that is the cornerstone of its educational mission. . . .

We take the Law School at its word that it would "like nothing better than to find a race-neutral admissions formula" and will terminate its race-conscious admissions program as soon as practicable. It has been 25 years since Justice Powell first approved the use of race to further an interest in student body diversity in the context of public higher education. Since that time, the number of minority applicants with high grades and test scores has indeed increased. We expect that 25 years from now, the use of racial preferences will no longer be necessary to further the interest approved today.

In summary, the Equal Protection Clause does not prohibit the Law School's narrowly tailored use of race in admissions decisions to further a compelling interest in obtaining the educational benefits that flow from a

diverse student body. Consequently, petitioner's statutory claims based on Title VI and 42 U.S.C. § 1981 also fail.

*Chief Justice Rehnquist, with whom  
Justices Scalia, Kennedy, and Thomas join, dissenting:*

I do not believe . . . that the University of Michigan Law School's means are narrowly tailored to the interest it asserts. The Law School claims it must take the steps it does to achieve a "critical mass" of underrepresented minority students. But its actual program bears no relation to this asserted goal. . . .

From 1995 through 2000, the Law School admitted between 1,130 and 1,310 students. Of those, between 13 and 19 were Native American, between 91 and 108 were African-Americans, and between 47 and 56 were Hispanic. If the Law School is admitting between 91 and 108 African-Americans in order to achieve "critical mass," thereby preventing African-American students from feeling "isolated or like spokespersons for their race," one would think that a number of the same order of magnitude would be necessary to accomplish the same purpose for Hispanics and Native Americans. . . .

Only when the "critical mass" label is discarded does a likely explanation for these numbers emerge. . . . [F]rom 1995 through 2000 the percentage of admitted applicants who were members of these minority groups closely tracked the percentage of individuals in the school's applicant pool who were from the same groups. . . . The tight correlation between the percentage of applicants and admittees of a given race, therefore, must result from careful race based planning by the Law School. . . . The Law School has offered no explanation for its actual admissions practices and, unexplained, we are bound to conclude that the Law School has managed its admissions program, not to achieve a "critical mass," but to extend offers of admission to members of selected minority groups in proportion to their statistical representation in the applicant pool. But this is precisely the type of racial balancing that the Court itself calls "patently unconstitutional."

*Justice Thomas, with whom  
Justice Scalia joins . . . , concurring in part and dissenting in part:*

The Law School adamantly disclaims any race-neutral alternative that would reduce "academic selectivity," which would in turn "require the Law School to become a very different institution, and to sacrifice a core part of its educational mission." In other words, the Law School seeks to improve marginally the education it offers without sacrificing too much of its exclusivity and elite status.

The proffered interest that the majority vindicates today, then, is not simply "diversity." Instead the Court upholds the use of racial discrimination

as a tool to advance the Law School's interest in offering a marginally superior education while maintaining an elite institution. Unless each constituent part of this state interest is of pressing public necessity, the Law School's use of race is unconstitutional. I find each of them to fall far short of this standard. . . .

Under the proper standard, there is no pressing public necessity in maintaining a public law school at all and, it follows, certainly not an elite law school. Likewise, marginal improvements in legal education do not qualify as a compelling state interest. . . .

The absence of any articulated legal principle supporting the majority's principal holding suggests another rationale. I believe what lies beneath the Court's decision today are the benighted notions that one can tell when racial discrimination benefits (rather than hurts) minority groups, and that racial discrimination is necessary to remedy general societal ills. . . .

I must contest the notion that the Law School's discrimination benefits those admitted as a result of it. . . . [N]owhere in any of the filings in this Court is any evidence that the purported "beneficiaries" of this racial discrimination prove themselves by performing at (or even near) the same level as those students who receive no preferences. . . . The Law School tantalizes unprepared students with the promise of a University of Michigan degree and all of the opportunities that it offers. These overmatched students take the bait, only to find that they cannot succeed in the cauldron of competition. . . . While these students may graduate with law degrees, there is no evidence that they have received a qualitatively better legal education (or become better lawyers) than if they had gone to a less "elite" law school for which they were better prepared. . . .

Beyond the harm the Law School's racial discrimination visits upon its test subjects, no social science has disproved the notion that this discrimination "engenders attitudes of superiority or, alternatively, provoke[s] resentment among those who believe that they have been wronged by the government's use of race." "These programs stamp minorities with a badge of inferiority and may cause them to develop dependencies or to adopt an attitude that they are 'entitled' to preferences."

It is uncontested that each year, the Law School admits a handful of blacks who would be admitted in the absence of racial discrimination. Who can differentiate between those who belong and those who do not? The majority of blacks are admitted to the Law School because of discrimination, and because of this policy all are tarred as undeserving. This problem of stigma does not depend on determinacy as to whether those stigmatized are actually the "beneficiaries" of racial discrimination. When blacks take positions in the highest places of government, industry, or academia, it is an open question today whether their skin color played a part in their advancement. . . .

**DISCUSSION QUESTIONS**

1. Reconstruct Justice O'Connor's main argument in the selection from her majority opinion in *Grutter*.
2. Reconstruct Chief Justice Rehnquist's main arguments in the selection from his dissenting opinion in *Grutter*. How could Justice O'Connor best respond? Is this response adequate? Why or why not?
3. What are Justice Thomas's main points in the selection from his dissenting opinion in *Grutter*? How could Justice O'Connor best respond? Is this response adequate? Why or why not?
4. How would you have decided the *Grutter* case? Why?

**GRATZ V. BOLLINGER**

(539 U.S. 244, 2003)

*Syllabus:*

Petitioners Gratz and Hamacher, both of whom are Michigan residents and Caucasian, applied for admission to the University of Michigan's (University) College of Literature, Science, and the Arts (LSA) in 1995 and 1997, respectively. Although the LSA considered Gratz to be well qualified and Hamacher to be within the qualified range, both were denied early admission and were ultimately denied admission. [The] University's Office of Undergraduate Admissions (OUA) . . . considers a number of factors in making admissions decisions, including high school grades, standardized test scores, high school quality, curriculum strength, geography, alumni relationships, leadership, and race. During all relevant periods, the University has considered African-Americans, Hispanics, and Native Americans to be "underrepresented minorities," and it is undisputed that the University admits virtually every qualified applicant from these groups. The current guidelines use a selection method under which every applicant from an underrepresented racial or ethnic minority group is automatically awarded 20 points of the 100 needed to guarantee admission. . . .

*Chief Justice Rehnquist delivered the opinion of the Court, in which Justices O'Connor, Scalia, Kennedy, and Thomas joined.*

We granted *certiorari* in this case to decide whether "the University of Michigan's use of racial preferences in undergraduate admissions violates the Equal Protection Clause of the Fourteenth Amendment, Title VI of the Civil Rights Act of 1964 . . . , or 42 U.S.C. § 1981. . . ." Because we find that the manner in which the University considers the race of applicants in its undergraduate admissions guidelines violates these constitutional and

statutory provisions, we reverse that portion of the District Court's decision upholding the guidelines. . . .

It is by now well established that "all racial classifications reviewable under the Equal Protection Clause must be strictly scrutinized." . . . To withstand our strict scrutiny analysis, respondents must demonstrate that the University's use of race in its current admission program employs "narrowly tailored measures that further compelling governmental interests." . . . We find that the University's policy, which automatically distributes 20 points, or one-fifth of the points needed to guarantee admission, to every single "underrepresented minority" applicant solely because of race, is not narrowly tailored to achieve the interest in educational diversity that respondents claim justifies their program.

In *Bakke*, Justice Powell . . . explained . . . that in his view it would be permissible for a university to employ an admissions program in which "race or ethnic background may be deemed a 'plus' in a particular applicant's file." . . . Justice Powell's opinion in *Bakke* emphasized the importance of considering each particular applicant as an individual, assessing all of the qualities that individual possesses, and in turn, evaluating that individual's ability to contribute to the unique setting of higher education. The admissions program Justice Powell described, however, did not contemplate that any single characteristic automatically ensured a specific and identifiable contribution to a university's diversity. . . . Instead, under the approach Justice Powell described, each characteristic of a particular applicant was to be considered in assessing the applicant's entire application.

The current LSA policy does not provide such individualized consideration. The LSA's policy automatically distributes 20 points to every single applicant from an "underrepresented minority" group, as defined by the University. The only consideration that accompanies this distribution of points is a factual review of an application to determine whether an individual is a member of one of these minority groups. Moreover, unlike Justice Powell's example, where the race of a "particular black applicant" could be considered without being decisive, the LSA's automatic distribution of 20 points has the effect of making "the factor of race . . . decisive" for virtually every minimally qualified underrepresented minority applicant.

Also instructive in our consideration of the LSA's system is the example provided in the description of the Harvard College Admissions Program, which Justice Powell both discussed in, and attached to, his opinion in *Bakke*. The example was included to "illustrate the kind of significance attached to race" under the Harvard College program. It provided as follows:

The Admissions Committee, with only a few places left to fill, might find itself forced to choose between A, the child of a successful black physician in an academic community with promise of superior academic performance, and B, a black who grew up in an inner-city ghetto of semi-literate parents whose academic achievement was lower but who had demonstrated energy and leadership as well as an apparently abiding interest in black power. If a good number

of black students much like A but few like B had already been admitted, the Committee might prefer B; and vice versa. If C, a white student with extraordinary artistic talent, were also seeking one of the remaining places, his unique quality might give him an edge over both A and B. Thus, the critical criteria are often individual qualities or experience *not dependent upon race but sometimes associated with it.* (emphasis added)

This example further demonstrates the problematic nature of the LSA's admissions system. Even if student C's "extraordinary artistic talent" rivaled that of Monet or Picasso, the applicant would receive, at most, five points under the LSA's system. At the same time, every single underrepresented minority applicant, including students A and B, would automatically receive 20 points for submitting an application. Clearly, the LSA's system does not offer applicants the individualized selection process described in Harvard's example. Instead of considering how the differing backgrounds, experiences, and characteristics of students A, B, and C might benefit the University, admissions counselors reviewing LSA applications would simply award both A and B 20 points because their applications indicate that they are African-American, and student C would receive up to 5 points for his "extraordinary talent."

Respondents emphasize the fact that the LSA has created the possibility of an applicant's file being flagged for individualized consideration by the ARC. We think that the flagging program only emphasizes the flaws of the University's system as a whole when compared to that described by Justice Powell. Again, students A, B, and C illustrate the point. First, student A would never be flagged. This is because, as the University has conceded, the effect of automatically awarding 20 points is that virtually every qualified underrepresented minority applicant is admitted. Student A, an applicant "with promise of superior academic performance," would certainly fit this description. Thus, the result of the automatic distribution of 20 points is that the University would never consider student A's individual background, experiences, and characteristics to assess his individual "potential contribution to diversity." Instead, every applicant like student A would simply be admitted.

It is possible that students B and C would be flagged and considered as individuals. This assumes that student B was not already admitted because of the automatic 20-point distribution, and that student C could muster at least 70 additional points. But the fact that the "review committee can look at the applications individually and ignore the points," once an application is flagged is of little comfort under our strict scrutiny analysis. The record does not reveal precisely how many applications are flagged for this individualized consideration, but it is undisputed that such consideration is the exception and not the rule in the operation of the LSA's admissions program. Additionally, this individualized review is only provided after admissions counselors automatically distribute the University's version of a "plus" that makes race a decisive factor for virtually every minimally qualified underrepresented minority applicant.

Respondents contend that “the volume of applications and the presentation of applicant information make it impractical for [LSA] to use the . . . admissions system” upheld by the Court today in *Grutter*. But the fact that the implementation of a program capable of providing individualized consideration might present administrative challenges does not render constitutional an otherwise problematic system. Nothing in Justice Powell’s opinion in *Bakke* signalled that a university may employ whatever means it desires to achieve the stated goal of diversity without regard to the limits imposed by our strict scrutiny analysis.

We conclude, therefore, that because the University’s use of race in its current freshman admissions policy is not narrowly tailored to achieve respondents’ asserted compelling interest in diversity, the admissions policy violates the Equal Protection Clause of the Fourteenth Amendment. . . .

*Justice Souter, with whom Justice Ginsburg joins as to Part II, dissenting:*

The record does not describe a system with a quota like the one struck down in *Bakke*, which “insulated” all nonminority candidates from competition from certain seats. . . . The plan here, in contrast, lets all applicants compete for all places and values an applicant’s offering for any place not only on grounds of race, but on grades, test scores, strength of high school, quality of course of study, residence, alumni relationships, leadership, personal character, socioeconomic disadvantage, athletic ability, and quality of a personal essay. A nonminority applicant who scores highly in these other categories can readily garner a selection index exceeding that of a minority applicant who gets the 20-point bonus.

Subject to one qualification to be taken up below, this scheme of considering, through the selection index system, all of the characteristics that the college thinks relevant to student diversity for every one of the student places to be filled fits Justice Powell’s description of a constitutionally acceptable program: one that considers “all pertinent elements of diversity in light of the particular qualifications of each applicant” and places each element “on the same footing for consideration, although not necessarily according them the same weight.” In the Court’s own words, “each characteristic of a particular applicant [is] considered in assessing the applicant’s entire application.” An unsuccessful nonminority applicant cannot complain that he was rejected “simply because he was not the right color”; an applicant who is rejected because “his combined qualifications . . . did not outweigh those of the other applicant” has been given an opportunity to compete with all other applicants.

The one qualification to this description of the admissions process is that membership in an underrepresented minority is given a weight of 20 points on the 150-point scale. On the face of things, however, this assignment of specific points does not set race apart from all other weighted considerations. Nonminority students may receive 20 points for athletic ability, socioeconomic disadvantage, attendance at a socioeconomically

disadvantaged or predominantly minority high school, or at the Provost's discretion; they may also receive 10 points for being residents of Michigan, 6 for residence in an underrepresented Michigan county, 5 for leadership and service, and so on.

The Court nonetheless finds fault with a scheme that "automatically" distributes 20 points to minority applicants because "the only consideration that accompanies this distribution of points is a factual review of an application to determine whether an individual is a member of one of these minority groups." The objection goes to the use of points to quantify and compare characteristics, or to the number of points awarded due to race, but on either reading the objection is mistaken.

The very nature of a college's permissible practice of awarding value to racial diversity means that race must be considered in a way that increases some applicants' chances for admission. Since college admission is not left entirely to inarticulate intuition, it is hard to see what is inappropriate in assigning some stated value to a relevant characteristic, whether it be reasoning ability, writing style, running speed, or minority race. Justice Powell's plus factors necessarily are assigned some values. The college simply does by a numbered scale what the [Michigan] law school accomplishes in its "holistic review"; the distinction does not imply that applicants to the undergraduate college are denied individualized consideration or a fair chance to compete on the basis of all the various merits their applications may disclose.

Nor is it possible to say that the 20 points convert race into a decisive factor comparable to reserving minority places as in *Bakke*. Of course we can conceive of a point system in which the "plus" factor given to minority applicants would be so extreme as to guarantee every minority applicant a higher rank than every nonminority applicant in the university's admissions system. But petitioners do not have a convincing argument that the freshman admissions system operates this way. The present record obviously shows that nonminority applicants may achieve higher selection point totals than minority applicants owing to characteristics other than race, and the fact that the university admits "virtually every qualified underrepresented minority applicant," may reflect nothing more than the likelihood that very few qualified minority applicants apply, as well as the possibility that self-selection results in a strong minority applicant pool. It suffices for me, as it did for the District Court, that there are no *Bakke*-like set-asides and that consideration of an applicant's whole spectrum of ability is no more ruled out by giving 20 points for race than by giving the same points for athletic ability or socioeconomic disadvantage. . . .

In contrast to the college's forthrightness in saying just what plus factor it gives for membership in an underrepresented minority, it is worth considering the character of one alternative thrown up as preferable, because supposedly not based on race. Drawing on admissions systems used at public

universities in California, Florida, and Texas, the United States contends that Michigan could get student diversity in satisfaction of its compelling interest by guaranteeing admission to a fixed percentage of the top students from each high school in Michigan.

While there is nothing unconstitutional about such a practice, it nonetheless suffers from a serious disadvantage.\* It is the disadvantage of deliberate obfuscation. The “percentage plans” are just as race conscious as the point scheme (and fairly so), but they get their racially diverse results without saying directly what they are doing or why they are doing it. In contrast, Michigan states its purpose directly and, if this were a doubtful case for me, I would be tempted to give Michigan an extra point of its own for its frankness. Equal protection cannot become an exercise in which the winners are the ones who hide the ball.

---

#### DISCUSSION QUESTIONS

1. Describe the main differences among the following admissions programs. Which of these differences are relevant to the constitutionality of these programs, in your opinion?
  - a. The admissions program at Michigan Law School
  - b. The admissions program at Michigan’s College of Literature, Science, and the Arts
  - c. The percentage plan used in California, Florida, and Texas
  - d. A lottery among all qualified candidates
2. In your opinion, would the admissions program at Michigan’s College of Literature, Science, and the Arts be unconstitutional if, instead of twenty points, it awarded only five points to members of underrepresented minority groups? What if it awarded fifty points? Does the number of points matter at all? Why or why not?
3. Much of the debate between Justices Rehnquist and Souter in *Gratz* concerns whether applicants receive “individualized consideration.” What exactly does this mean? Why is it important in this context?
4. What are the implications of *Grutter* and *Gratz* for admissions policies in private colleges and universities? For preferential treatment in hiring? In government contracts?
5. Find out what kinds of preferential treatment programs exist in your own school or town, and then argue either that these programs are constitutional or that they are not.

\* Of course it might be pointless in the State of Michigan, where minorities are a much smaller fraction of the population than in California, Florida, or Texas.

## BURDEN OF PROOF

A remarkable feature of the line of cases from *Plessy* through *Brown* to *Bakke*, *Grutter*, and *Gratz* is the extent to which interpretations of the equal protection clause turn on the matter of *burden of proof*. Under the rational-relation test that governed *Plessy*, the state bears a light burden when it is asked to show that its actions do not conflict with the equal protection clause. In contrast, the strict-scrutiny test that governed *Brown*, that Powell applied in *Bakke*, and that the Court applies in *Grutter* and *Gratz* places a very heavy burden on the state to justify any use of a suspect classification or any interference with fundamental rights.

It may seem peculiar that an important legal decision can turn on such a technical and procedural matter as burden of proof. But the question of burden of proof often plays a decisive role in a legal decision, so it is worth knowing something about it.

The two basic questions concerning burden of proof are (1) *who* bears this burden and (2) how *heavy* is the burden. In our system of criminal justice, the rules governing burden of proof are fairly straightforward. The state has the burden of establishing the guilt of the accused. The defendant has no obligation to establish his or her innocence. That is what is meant by saying that the defendant is *innocent until proven guilty*. The burden of proof is also very heavy on the state in criminal procedures, for it must show *beyond a reasonable doubt* that the accused is guilty. If the prosecution shows only that it is more likely than not that the accused has committed a crime, then the jury should vote for acquittal.

Turning to civil law, there is no simple way to explain burden of proof. Very roughly, the plaintiff (the one who brings the suit) has an initial burden to establish a *prima facie* case—that is, a case that is strong enough that it needs to be rebutted—on behalf of his or her complaint. The burden then shifts to the respondent (the one against whom the suit is being brought) to answer these claims. The burden may then shift back and forth depending on the nature of the procedure. Provided that both sides have met their legally required burdens of proof, the case is then decided on the basis of the *preponderance of evidence*; that is, the judge or jury decides which side has made the stronger case.

Burden of proof is primarily a legal notion, but it is sometimes used, often loosely, outside the law. The notion of burden of proof is needed within the law because law cases are adversarial and the court has to come to a decision. Outside the law, people have a general burden to *have* good reasons for what they say. That is the second part of Grice's rule of Quality (described in Chapter 2). More specifically, people have a burden to be able to *present* some reasons when they make accusations or statements that run counter to common opinion.

The important thing to see is that you cannot establish the truth of something through an appeal to the burden of proof. The following argument is perfectly weird:

There is life in other parts of the universe, because you can't prove otherwise.

Of course, no one can prove that there is *not* life elsewhere in the universe, but this has no tendency to show that there *is*. Attempts to prove the truth of something through appeals to burden of proof—often called arguments from ignorance—are another example of a fallacy of relevance. (See Chapter 15.)

Nonetheless, the importance of burden of proof in the law does give force to another kind of argument. In a criminal case, the following argument would be perfectly fine:

The defendant ought to be found not guilty, because the prosecution has not proven beyond a reasonable doubt that she is guilty.

This argument would also be a fallacy of relevance if the burden of proof were not so important. But the relevant burden of proof makes this argument strong in a court of law. Who bears the burden of proof and how heavy the burden is, determine which legal arguments work. Consider the following argument:

This law uses a suspect classification, and the state has not shown that it serves any compelling purpose, so the law is unconstitutional.

This argument is strong if the strict-scrutiny interpretation is accepted (assuming the premises are true). But this argument fails if a weaker burden of proof is required, as in the rational-relation test. When one chooses between interpretations of the equal protection clause and between different burdens of proof, one also chooses which arguments will have force in courts of law. This is another example of a general phenomenon that has been stressed throughout this book—that background assumptions can determine whether an argument is any good.

#### DISCUSSION QUESTION

Try to formulate general rules governing who has the burden of proof when people disagree. Be sure to consider a variety of areas, such as science, religion, philosophy, morality, and personal life.

## MORAL REASONING

*Many acts—such as lying to a friend—are not illegal, but they still seem immoral. Thus, even if such an act is legally permitted, this does not show that it is morally permitted. That is a separate issue, and it is one that many people care deeply about, because they want to do what is moral and avoid doing what is immoral. But how can we show that an act is moral or immoral? One kind of argument will not do. We can often show that an act is illegal by citing official pronouncements by judges and legislators, such as precedents and statutes. In contrast, morality is not decided by any official. There are no authoritative books in which we can look up whether a certain act is immoral without asking whether that book is correct. This affects the nature of moral arguments and the criteria for evaluating them. We cannot appeal to any documents or officials to justify our moral beliefs, so moral beliefs must be based on something else. The kinds of arguments that can be used to justify moral beliefs will be the topic of this chapter.*

## MORAL DISAGREEMENTS

---

People often disagree on moral questions. When these disagreements arise, it is often difficult—and sometimes impossible—to resolve them. At times these disagreements turn on questions of fact. If one person thinks that an action will have a particular consequence, and another thinks that it will not, they might well disagree on the moral worth of that action. For example, those who have defended the United States' decision to drop atomic bombs on Japan have often claimed that it was the only way to end the war quickly without creating a great number of casualties on both sides. Many critics of this decision have denied this factual claim.

Moral disagreements can also arise from disagreements about moral principles. To many people, it is immoral to have sex outside marriage. To others, it is immoral to interfere with such acts.

Despite such disagreements, it is surprising how much agreement there is on general moral principles. In our society, most people accept a great many moral principles as a matter of course. If a policy has no other consequence but to produce widespread misery, it is rejected out of hand.

We share a conception of justice that includes, among other things, equality of opportunity and equality before the law. Most people also have a conception of human dignity: A human being is not a thing to be used and disposed of for personal advantage.

With all this agreement, how does moral disagreement arise at all? One answer is that in certain circumstances, our moral principles *conflict* with one another, and people are inclined to resolve these conflicts in different ways. People often agree on principles about welfare, justice, and human dignity, and yet, by weighing these principles differently or seeing the situation in a different light, they arrive at opposing moral conclusions.

Another kind of moral disagreement concerns the *range* or *scope* of moral principles. Even if everyone agrees that death and suffering are bad, they often disagree about *whose* death and suffering count. With few exceptions, it is thought to be wrong to inflict death and suffering on human beings. Most people have a similar attitude toward their pet dogs or cats. Some, however, go further and claim that it is also immoral to kill any animals—including cows, chickens, and fish—just to produce tasty food for humans.

The hardest problems combine issues of *range* with *conflicts of principles*. It is a disagreement of this complex kind that we will focus on in this chapter. The problem is abortion. The main issues are (1) whether fetuses lie within the range of a standard moral principle against killing, and (2) how to resolve conflicts between the principles that protect fetuses and other principles concerning, for example, human welfare and a woman's control over her body.

## THE PROBLEM OF ABORTION

---

When faced with a moral problem, it often seems clear what the problem is, but this assumption can be mistaken. Sometimes a problem is formulated so vaguely that there is no way even to begin to solve it. People can argue for hours or even years without realizing that they are really talking about different things.

To clarify a moral problem, the first step is to specify precisely what is being judged—which action or kind of action is at issue. In the problem of abortion, the first step is to specify exactly what counts as an abortion. It is common to define abortion as the termination of a pregnancy. This includes spontaneous abortions or miscarriages, but these raise no moral problems because they are not the result of human action. Furthermore, the moral problem of abortion arises only when the death of the fetus is an expected consequence of terminating the pregnancy. To focus on these problematic cases, from now on we will take “abortion” to mean the intentional termination of a pregnancy with the expected consequence that the fetus dies as a result.

After the class of actions is picked out, we need to determine what is being asked about this class of actions—what kind of moral judgment is at stake. It is one thing to ask whether abortion is *morally wrong*, and another thing to ask whether abortion *should be illegal*. These are both moral questions (because the second asks what the law should be and not what it is), but they can be answered differently. It is not uncommon for people to claim that abortion is morally wrong but should not be made illegal, because it is a matter of personal, not public, morality. It is also important to distinguish the question of whether abortion is or is not morally *wrong* from the separate question of whether abortion is or is not *good*. People who deny that abortion is morally wrong do not hold that abortion is a positive good. They do not, for example, recommend that people get pregnant so that they can have abortions. So, from now on, we will focus on the issue of the moral wrongness of abortion.

### THE “PRO-LIFE” ARGUMENT

We can begin to understand this problem if we reconstruct the main argument against abortion, using the method sketched in Chapter 5. Most opponents of abortion call themselves “pro-life” and base their position on an appeal to a moral principle involving the “right to life.” Of course, most opponents of abortion are not opposed to killing weeds, germs, or even fish. What they have in mind, then, is probably a principle such as this:

It is always morally wrong to kill a human being.

This principle by itself does not rule out abortion. To reach this conclusion, we need further premises of the following kind:

Abortion involves killing a human fetus.

A human fetus is a human being.

With these premises, the anti-abortion argument will have the following form:

- (1) It is always morally wrong to kill a human being.
  - (2) Abortion involves killing a human fetus.
  - (3) A human fetus is a human being.
- ∴ (4) Abortion is always morally wrong.

This argument is valid and reasonably charitable, so we have completed the first stage of reconstruction.

We next ask if the premises of this argument are true. The second premise is not controversial, given our definition of abortion; but the third premise raises many problems. Much of the debate concerning abortion turns on the question of whether a fetus is a human being. We will examine this question later on. For now, we will assume for the sake of argument that a fetus is a human being. That leaves only the first premise.

Some people—for example, strong pacifists—accept premise (1), but most people who adopt strong anti-abortion positions do not. This comes out in the following way. Many of those who oppose abortion are in favor of the death penalty for certain crimes. Therefore, they do not accept the general principle that it is always wrong to take a human life. What they need, then, is a principle that allows taking a human life in some instances but not in others. In an effort to achieve this, those who oppose abortion could reformulate the first premise in these words:

(1\*) It is always morally wrong to kill an *innocent* human being.

Here the word “innocent” allows an exception for the death penalty being imposed on those who are found guilty of certain crimes. Of course, if we simply stick this premise into the previous argument, the result is invalid:

(1\*) It is always morally wrong to kill an innocent human being.

(2) Abortion involves killing a human fetus.

(3) A human fetus is a human being.

(4) Abortion is always morally wrong.

To make it valid, we need to add a new premise, so that the whole argument looks like this:

(1\*) It is always morally wrong to kill an innocent human being.

(2) Abortion involves killing a human fetus.

(3) A human fetus is a human being.

(4\*) A human fetus is innocent.

. . . (5) Abortion is always morally wrong.

Even stated this way, however, the first premise seems to admit of counterexamples. If someone’s life is threatened by a madman, it is generally thought that the person has the right to use whatever means are necessary against the madman to prevent being killed. This may include killing the madman, even though the insane are usually thought to be morally innocent of their deeds. If so, the moral principle must be modified again, and then we get something like this:

(1\*\*) It is always morally wrong to kill an innocent human being *except in self-defense.*

It is still possible to find difficulties with this principle that will lead some to add further modifications or clarifications. Children, for example, are often the innocent victims of bombing raids, yet the raids are often thought to be justified, because these deaths are not intended, even though they are foreseeable. At this point it is common to modify the principle again by including a reference to intentions. We shall not, however, pursue this complex line of reasoning here.<sup>1</sup>

We have arrived, then, at a principle that seems to make sense out of a position that is against abortion but in favor of the death penalty and self-defense. With these modifications included, the argument now looks like this:

- (1\*\*) It is always morally wrong to kill an innocent human being except in self-defense.
  - (2) Abortion involves killing a human fetus.
  - (3) A human fetus is a human being.
  - (4\*) A human fetus is innocent.
- ∴ (5) Abortion is always morally wrong.

Again, however, by making the premises more plausible, we created a new problem: The argument is invalid as it stands, since the qualification “except in self-defense” is missing from the conclusion. The proper conclusion of the argument should be:

- (5\*) Abortion is always morally wrong except in self-defense.

Rewriting the conclusion in this way has an important consequence: The argument no longer leads to a conclusion that abortion is *always* wrong. This qualified conclusion could permit abortion in those cases in which it is needed to defend the life of the pregnant woman who bears the fetus. In fact, this is the position that many people who are generally opposed to abortion adopt: Abortion is wrong except in those cases in which it is necessary to save the life of the mother. Although this does not lay down an absolute prohibition, it is still a strong anti-abortion position, since it would condone abortion in only a few exceptional cases.

### **“PRO-CHOICE” RESPONSES**

We can now examine the way in which those who adopt the liberal or “pro-choice” position will respond to the conservative or “pro-life” argument as it has just been spelled out. The second premise should not be a subject for controversy, given our definition of abortion. Nor does it seem likely that the fourth premise will be attacked on the ground that the fetus is not innocent. How could a fetus be guilty of anything?

This leaves three strategies for the liberal: (1) Further modify the moral principle in the first premise to allow more exceptions. (2) Deny the third premise—that the fetus is a human being. (3) Oppose this conservative argument with a different argument based on a different moral principle.

*FURTHER MODIFICATIONS.* Even if it is agreed that abortion is justified when it saves the mother’s life, we still need to ask whether this is the only exception or whether abortion is justified in other cases as well. Many pro-life conservatives admit that abortion is also justified when the pregnancy results from rape or incest. It is not easy to see how to modify the moral principle against killing to allow an exception in cases of rape and incest, so

this exception is controversial. We will return to this issue later. But even if exceptions are made both for life-threatening pregnancies and for pregnancies due to rape and incest, the range of morally permissible abortions will still be very small.

Pro-choice liberals can, however, argue for a wider range of morally permissible abortions by extending the self-defense exception. It can be argued that a woman has a right to defend not only her life but also her physical and psychological well-being. Liberals can also argue that the exception of rape shows that abortion is allowed when the woman is not responsible for her pregnancy, and this might include cases in which the woman tried to prevent pregnancy by using contraceptives. Granting exceptions of this kind does not provide the basis for an absolute right to an abortion, but it does move things away from a “pro-life” position in the direction of a “pro-choice” position.

*THE STATUS OF THE FETUS.* So far we have assumed for the sake of argument that a human fetus is a human being. But pro-choice liberals often deny this premise. It may seem hard to deny that a human fetus is human. After all, it is not an aardvark. Liberals, however, claim that the real issue is not about biological species. The real issue is whether a human fetus is covered by the moral principle against killing, and whether it is protected to the same extent as an adult human. Anything that is protected to this extent is said to have a “right to life” and will be called a “person.” The issue, then, is whether a human fetus is a person. If a fetus is a person, the burden of proof is on those who defend abortion to show why the moral principle against killing should be set aside or modified. If a fetus is *not* a person, this moral principle cannot show that there is anything wrong with abortion for any reason—with what is called “abortion on demand.”

Any argument that a fetus either is a person or is not a person must proceed from some idea of which properties make something a person—which properties warrant the protection of moral principles. To argue that a fetus is not a person, pro-choice liberals need to find some feature that fetuses lack and that is necessary for personhood. In response, pro-life conservatives need to find some feature that fetuses have and that is sufficient for personhood.

Many conditions of personhood have been suggested. This list is not complete:

Genetic code (which determines biological species)

Ensoulment (when a soul enters the body)

Brain activity (first detected around eight weeks)

Sentience (capacity to feel pain and pleasure)

Viability (when the fetus can survive outside the womb)

Rationality (and other related capacities)

Pro-life conservatives usually emphasize tests such as genetic code (which is formed at conception) or ensoulment (which is supposed to occur at or shortly after conception). In contrast, pro-choice liberals usually employ tests such as viability (which is reached during the second trimester) or rationality (which comes sometime after birth, depending on what counts as rationality—ability to choose and plan, self-consciousness, and so on). Thus, the personhood of fetuses during the first trimester is usually asserted by conservatives and denied by liberals on this issue.

How can we determine whether a feature is necessary or sufficient for personhood? We can start by rejecting any test of personhood that leads to *implausible* results. Pro-life conservatives argue that rationality is not necessary for personhood, because, whatever rationality is, newborn babies and severely retarded adults are not rational, but it is still morally wrong to kill them. Other tests of personhood are ruled out because they do not seem *important* enough. Pro-choice liberals argue that a certain genetic code is not sufficient to make something a person, because there is no reason to favor one genetic code over another except that it later produces other important features, such as rationality. It is also common to rule out a test of personhood if we cannot *know* when the test is passed. For example, many people reject ensoulment as a criterion of personhood, because they see no way to tell when, if ever, a fetus has a soul. And tests of personhood are also often rejected if they depend on factors that are *extraneous*. Conservatives often argue that viability cannot be a test of personhood, because the point when a fetus can survive outside the womb depends on what technology happens to be available to doctors at the time.

In addition to features that fetuses have when they are fetuses, they also seem to have the *potential* to develop many more, including rationality. Opponents of abortion often use this premise to argue that fetuses are persons and have a right to life. The first problem with this argument is that it seems to assume that something has a right if it has the potential to come to have that right. But this is clearly too strong. A three-year-old child does not have the right to vote even though it has the potential to develop into someone who will have the right to vote. Furthermore, the notion of potential is not clear. If the fetus has the potential, why do the egg and sperm not have it? This does not refute potentiality as a test of personhood, but much more must be done to show what potentiality is and why it is sufficient to make something a person even before the potential is realized.

All these positions on personhood are controversial, and many people feel uncertain about which is the correct one. A major issue in many moral problems is how to deal with uncertainties such as this. One reaction is a position called “gradualism.”<sup>2</sup> We have assumed so far that the fetus either has a full right to life or has no right to life at all, but rights sometimes come in varying strengths. Gradualists claim that a fetus slowly develops a right to life that is at first very weak. As pregnancy progresses, this right gets stronger, so it takes more to justify abortion. Late abortions still might be

permitted, but only in extreme circumstances. This position is still vague, but it is attractive to some people who want to avoid placing too much emphasis on any single point in fetal development.

Uncertainty is also exploited in other ways. We already discussed slippery-slope arguments in Chapter 13. Another way to exploit uncertainty is to put the burden on the other side to produce a reason for drawing a line at some point. For example, Ronald Reagan said, "Anyone who doesn't feel sure whether we are talking about a second human life should clearly give life the benefit of the doubt." However, the same kind of argument is also available to defenders of abortion: Since we are not sure whether the fetus is a person, but we are sure that the pregnant woman has rights over her body, we should give the benefit of the doubt to the pregnant woman. We should always suspect that there is something wrong with an argument that can be used equally well in opposing directions.

*CONFLICTING PRINCIPLES.* A third kind of pro-choice response is to invoke another principle, which conflicts with the pro-life principle against killing. Pro-choice liberals often emphasize two such principles: one about the rights of the pregnant woman to control her own body and another about overall human welfare. We will focus for now on human welfare.

Defenders of abortion often argue that abortion can sometimes be justified in terms of the welfare of the woman who bears the fetus, or in terms of the welfare of the family into which it will be born, or even in terms of the welfare of the child itself (if it were to be born with a severe disability or into an impoverished situation). This argument, when spelled out, looks like this:

- (1) An action that best increases overall human welfare is not morally wrong.
  - (2) Abortion is sometimes the best way of increasing overall human welfare.
- ∴ (3) Abortion is sometimes not morally wrong.

What are we to say about this argument? It seems valid in form, so we can turn to the premises themselves and ask whether they are acceptable. The first (and leading) premise of the argument is subject to two immediate criticisms. First, it is vague. Probably what a person who uses this kind of argument has in mind by speaking of human welfare is a certain level of material and psychological well-being. Of course, this is still vague, but it is clear enough to make the premise a target of the second, more important, criticism: Although maximizing human welfare may, in general, be a good thing, it is not the only relevant consideration in deciding how to act. For example, it might be true that our society would be much more prosperous on the whole if 10 percent of the population were designated slaves who would do all the menial work. Yet, even if a society could be made generally happy in this way, most people would reject such a system on the grounds

that it is unfair to the slave class. (See the discussion of counterexamples in Chapter 17.) For reasons of this kind, most people would modify premise (1) in the following way:

- (1\*) An action that best increases human welfare is not morally wrong,  
provided that it is not unfair.

But if the first premise is modified in this way, then the entire argument must be restated to reflect this revision. It will now look like this:

- (1\*) An action that best increases human welfare is not morally wrong,  
provided that it is not unfair.  
(2\*) Abortion is sometimes an action that best increases human welfare.  
. . . (3\*) Abortion is sometimes not morally wrong, provided that it is not  
unfair.

It should be obvious how opponents of abortion will reply to this argument. They will maintain that abortion almost always involves unfairness—namely, to the fetus—so abortion is still wrong in almost all cases, as the pro-life argument claimed. Once more we have encountered a standard situation: Given a strong premise (premise 1), it is possible to derive a particular conclusion, but this strong premise is subject to criticism and therefore must be modified. When the premise is modified as in 1\*, it no longer supports the original conclusion that the person presenting the argument wishes to establish.

The argument does not stop here. A defender of abortion might reply in a number of ways. Some theory of fairness might be developed to argue that many abortions are not unfair to the fetus, because the fetus has no right to use the pregnant woman's body. The burden of the argument may shift to the question of whether or not a human fetus is a person and therefore possessed of a right to fair treatment. It might also be argued that questions of human welfare are sometimes more important than issues of fairness. During war and some emergencies, for example, members of a certain segment of the population are called on to risk their lives for the good of the whole in ways that might seem unfair to them.

When the argument is put on this new basis, the question becomes this: Are there circumstances in which matters of welfare become so urgent that the rights of the fetus (assuming the fetus has rights) are overridden? The obvious case in which this might happen is when the life of the bearer of the fetus is plainly threatened. For many conservatives on abortion, abortion is permitted in such cases. Some who hold a pro-choice position will maintain that severe psychological, financial, or personal losses to the pregnant woman may also take precedence over the life of the fetus. Furthermore, if not aborted, many fetuses would live in very deprived circumstances, and some would not develop very far or live very long, because they have deadly diseases, such as Tay-Sachs. How severe must these losses, deprivations, and diseases be? From our previous discussion of slippery-slope

arguments in Chapter 13, we know that we should not expect any sharp lines here. Indeed, people will tend to be spread out in their opinions along a continuum ranging from a belief in complete prohibition to no prohibition.

## **ANALOGICAL REASONING IN ETHICS**

---

Using the method for reconstructing arguments, we now have a fairly clear idea of the main options on the abortion issue. But understanding the structure of the debate—though essential for dealing with it intelligently—does not settle it. If the reasons on all sides are fully spelled out and disagreement remains, what is to be done?

At this stage, those who do not simply turn to abuse often appeal to *analogical arguments*. The point of an analogical argument is to reach a conclusion in a controversial case by comparing it to a similar situation in which it is clearer what is right or wrong. In fact, a great deal of ethical reasoning uses such analogies. We have already seen one simple analogy between an abortion to save the life of the mother and self-defense against an insane person. To get a better idea of how analogical reasoning works in ethics, we will concentrate on a more complex analogy, which raises the issue of whether abortion is morally permissible in cases of pregnancy due to rape.

A classic analogical argument is given by Judith Jarvis Thomson in “A Defense of Abortion.” Thomson grants for the sake of argument that a fetus is a person. Then she tells the following story:

You wake up in the morning and find yourself back to back in bed with an unconscious violinist. A famous unconscious violinist. He has been found to have a fatal kidney ailment, and the Society of Music Lovers has canvassed all the available medical records and found that you alone have the right blood type to help. They have therefore kidnapped you, and last night the violinist’s circulatory system was plugged into yours, so that your kidneys can be used to extract poisons from his blood as well as your own. The director of the hospital now tells you, “Look, we’re sorry the Society of Music Lovers did this to you—we would never have permitted it if we had known. But still, they did it, and the violinist now is plugged into you. To unplug you would be to kill him. But never mind, it’s only for nine months. By then he will have recovered from his ailment, and can safely be unplugged from you.”<sup>3</sup>

Thomson claims that it is not wrong for you to unplug yourself from the violinist in this situation, and most people seem to agree with this judgment. By analogy, abortion after rape is not wrong either, or so she says.

The basic assumption of this analogical argument is that we should not make different moral judgments in cases that do not differ. More positively:

- (1) If two actions are similar in all morally relevant respects, and if one of the acts is not morally wrong, then the other act is also not morally wrong.

Now we can apply this principle to Thomson's story:

- (2) It is not morally wrong for you to unplug the violinist in Thomson's example.
- (3) To unplug the violinist and to abort a pregnancy due to rape are similar in all morally relevant respects.
- ∴ (4) It is not morally wrong for a woman to abort a pregnancy due to rape.

This argument is valid, so, following the normal procedure, we can ask whether the premises are true. The first premise seems plausible, and it is accepted in most moral theories. Most people also accept the second premise. Consequently, the discussion usually focuses on the third premise—on the similarities and differences between Thomson's story and abortion in a pregnancy due to rape.

First, consider these similarities between Thomson's story and abortion after rape:

1. Both the fetus and the violinist are on or near the Earth.
2. Kidnapping is immoral and illegal, like rape.
3. The hospital stay lasts nine months, like pregnancy.
4. The violinist is innocent and a human being, like the fetus (given our present assumption).
5. Unplugging the violinist is supposed to be killing, like an abortion.

Now here are some differences between the situations:

6. The fetus cannot play the violin, but the violinist can.
7. The person who is plugged into the violinist might not be female.
8. The person who is plugged into the violinist cannot leave the hospital room, but pregnant women can still move around, even if they have some difficulty.
9. Abortion involves killing, but unplugging the violinist is merely refusing to save.

It is obvious that some of the similarities and differences are *not* relevant. It does not matter whether killing occurs near the Earth. Killing is usually wrong, even on the starship *Enterprise*. It is also accepted that differences in musical talent and in sex cannot justify killing. The other similarities and differences on our list *do* seem important. They each concern harm and responsibility, matters that must be considered in reaching a moral judgment about these actions. The force of Thomson's analogical argument is that the very features that lead us to conclude that it would not be wrong to unplug the violinist are also found in the case of pregnancy due to rape. Furthermore, there are no relevant differences that are important enough to override the significance of these similarities. These considerations, if correct, provide a reason for treating the two cases in the same way. If we agree, as Thomson

thinks we will, that it is not wrong to unplug the violinist, then we have a reason to conclude that abortion after rape is not wrong either.

Responses to Thomson's argument have largely turned on emphasizing the differences between the two situations. Many critics claim that Thomson's argument fails because abortion involves *killing*, whereas unplugging does not. If you stay plugged to the violinist, this will save the violinist, so to unplug yourself is to *fail to save* the violinist. But critics deny that to unplug yourself from the violinist is the same as to *kill* the violinist, or to take the violinist's life. They argue that there is a crucial difference between killing and failing to save, because a negative duty not to kill is much stronger than any positive duty to save another person's life.

To determine whether unplugging the violinist is more like acts of killing or more like other acts of refusing to save, we might consider more analogies. Thomson also introduces additional analogies that seem more like abortions in which the pregnancy is not due to rape. In the end, our sense of which features seem most important will determine how we evaluate all such analogical arguments. The analogies bring certain features to our attention, but we have to decide which features are important, and how important they are.

## WEIGHING FACTORS

Our discussion has brought us to the following point: Disagreements concerning abortion in general cannot be reduced to a yes–no dispute. Most opponents of abortion acknowledge that abortion is permissible in some (though very few) cases. Most defenders of abortion admit that there are some (though not restrictively many) limitations on when abortion is permissible. Where people place themselves on this continuum does not depend on any simple acceptance of one argument over another, but instead on the weight they give certain factors. To what extent does a fetus have rights? The pro-life position we examined earlier grants the fetus a full (or close to full) right to life. The pro-choice position usually grants few or no rights to the fetus. In what areas do questions of welfare override certain individual rights? The conservative in this matter usually restricts this to those cases in which the very life of the mother is plainly threatened. As the position on abortion becomes more liberal, the more extensive becomes the range of cases in which the rights, if any, of the fetus are set aside in favor of the rights of the bearer of the fetus. Where a particular person strikes this balance is not only a function of basic moral beliefs but also a function of different weights assigned to them.

How can one deal with such bedrock disagreements? The first thing to see is that logic alone will not settle them. Starting from a certain conception of persons, it is possible to argue coherently for a pro-choice view on abortion; starting from another point of view, it is possible to argue coherently for a pro-life view on abortion. The next important thing to see is that it is

possible to *understand* an opposing view—that is, get a genuine feeling for its inner workings—even if you disagree with it completely. Logical analysis might show that particular arguments are unsound or have unnoticed and unwanted implications. This might force clarification and modification. But the most important service that logical analysis can perform is to lay bare the fundamental principles that lie beneath surface disagreements. Analysis will sometimes show that these disagreements are fundamental and perhaps irreconcilable. Dealing with such irreconcilable differences in a humane way is one of the fundamental tasks of a society dedicated to freedom and a wide range of civil liberties.

#### DISCUSSION QUESTIONS

1. Reconstruct and evaluate the arguments against abortion that are stated or suggested in the following short passages from Ronald Reagan, *Abortion and the Conscience of the Nation* (New York: Thomas Nelson, 1984). Be sure to specify the exact conclusion and spell out important suppressed premises. How would an opponent best respond to each argument?
  - a. "We cannot diminish the value of one category of human life—the unborn—without diminishing the value of all human life" (18).
  - b. "I have often said that when we talk about abortion, we are talking about two lives—the life of the mother and the life of the unborn child. Why else do we call a pregnant woman a mother?" (21).
  - c. "I have also said that anyone who doesn't feel sure whether we are talking about a second human life should surely give life the benefit of the doubt. If you don't know whether a body is alive or dead, you would never bury it. I think this consideration itself should be enough for all of us to insist on protecting the unborn" (21).
  - d. "Medical practice confirms at every step the correctness of these moral sensibilities. Modern medicine treats the unborn child as a patient. Medical pioneers have made great breakthroughs in treating the unborn—for genetic problems, vitamin deficiencies, irregular heart rhythms, and other medical conditions" (21–22).
  - e. "I am convinced that Americans do not want to play God with the value of human life. It is not for us to decide who is worthy to live and who is not" (30).
  - f. "Malcolm Muggeridge, the English writer, goes right to the heart of the matter: 'Either life is always and in all circumstances sacred, or intrinsically of no account; it is inconceivable that it should be in some cases the one and in some the other'" (34).
2. Reconstruct and evaluate the arguments in defense of abortion that are stated or suggested in the following short passages from Mary Gordon, "A Moral Choice," *Atlantic Monthly*, April 1990, 78–84. Be sure to specify the exact conclusion and spell out important suppressed premises. How would an opponent best respond to each argument?

(continued)

- a. “Common sense, experience, and linguistic usage point clearly to the fact that we habitually consider, for example, a seven-week-old fetus to be different from a seven-month-old one. . . . We have different language for the experience of the involuntary expulsion of the fetus from the womb depending upon the point of gestation at which the expulsion occurs. If it occurs early in the pregnancy, we call it a miscarriage; if late, we call it a stillbirth” (80).
- b. “Our ritual and religious practices underscore the fact that we make distinctions among fetuses. If a woman took the bloody matter—indistinguishable from a heavy period—of an early miscarriage and insisted upon putting it in a tiny coffin and marking its grave, we would have serious concerns about her mental health. By the same token, we would feel squeamish about flushing a seven-month-old fetus down the toilet—something we would normally do with an early miscarriage. There are no prayers for the matter of a miscarriage, nor do we feel there should be. Even a Catholic priest would not baptize the issue of an early miscarriage” (80).
- c. “We must make decisions on abortion based on an understanding of how people really do live. We must be able to say that poverty is worse than not being poor, that having dignified and meaningful work is better than working in conditions of degradation, that raising a child one loves and has desired is better than raising a child in resentment and rage, that it is better for a twelve-year-old not to endure the trauma of having a child when she is herself a child” (81–82).
- d. “It is possible for a woman to have a sexual life unriddled by fear only if she can be confident that she need not pay for a failure of technology or judgment (and who among us has never once been swept away in the heat of a sexual moment?) by taking upon herself the crushing burden of unchosen motherhood” (82).
- e. “There are some undeniable bad consequences of a woman’s being forced to bear a child against her will. First is the trauma of going through a pregnancy and giving birth to a child who is not desired, a trauma more long-lasting than that experienced by some (only some) women who experience an early abortion. The grief of giving up a child at its birth—and at nine months it is a child one has felt move inside one’s body—is underestimated both by anti-choice partisans and by those for whom access to adoptable children is important. This grief should not be forced on any woman—or, indeed, encouraged by public policy” (84).
- f. “We must be realistic about the impact on society of millions of unwanted children in an overpopulated world” (84).
- g. “Making abortion illegal will result in the deaths of women, as it has always done. Is our historical memory so short that none of us remember aunts, sisters, friends, or mothers who were killed or rendered sterile by septic abortions? . . . Can anyone genuinely say that it would be a moral good for us as a society to return to those conditions?” (84).

## ABORTION<sup>4</sup>

by Mary Anne Warren

### I INTRODUCTION

Do women have the right to abort unwanted pregnancies? Or is the state entitled (or perhaps ethically required) to prohibit deliberate abortion? Should some abortions be permitted and others not? Does the proper legal status of abortion follow directly from its moral status? Or should abortion be legal, even if it is sometimes or always morally wrong?

Such questions have aroused intense debate during the past two decades. Interestingly enough, in most of the industrialized world abortion was not a criminal offence until a series of anti-abortion laws were passed during the second half of the nineteenth century. At that time, proponents of the prohibition of abortion generally stressed the medical dangers of abortion. It was also sometimes argued that fetuses are human beings from conception onward, and that deliberate abortion is therefore a form of homicide. Now that improved techniques have made properly performed abortions much safer than childbirth, the medical argument has lost whatever force it may once have had. Consequently, the focus of anti-abortion arguments has shifted from the physical safety of women to the moral value of fetal life.

Advocates of women's right to choose abortion have responded to the anti-abortion argument in several ways. I shall examine three lines of argument for the pro-choice view: (1) that abortion should be permitted, because the prohibition of abortion leads to highly undesirable consequences; (2) that women have a moral right to choose abortion; and (3) that fetuses are not yet persons and thus do not yet have a substantial right to life.

### II CONSEQUENTIALIST ARGUMENTS FOR ABORTION

If actions are to be morally evaluated by their consequences, then a strong case can be made that the prohibition of abortion is wrong. Throughout history women have paid a terrible price for the absence of safe and legal contraception and abortion. Forced to bear many children, at excessively short intervals, they were often physically debilitated and died young — a common fate in most pre-twentieth-century societies and much of the Third World today. Involuntary childbearing aggravates poverty, increases infant and child death rates, and places severe strains upon the resources of families and states.

Improved methods of contraception have somewhat alleviated these problems. Yet no form of contraception is 100 per cent effective. Moreover, many women lack access to contraception, e.g. because they cannot afford it, or it is unavailable where they live, or unavailable to minors without parental permission. In most of the world, paid work has become an economic necessity

for many women, married or single. Women who must earn have an acute need to control their fertility. Without that control, they often find it impossible to obtain the education necessary for any but the most marginal employment, or impossible to combine the responsibilities of childrearing and paid labour. This is as true in socialist as in capitalist economies, since in both economic systems women must contend with the double responsibility of paid and domestic work.

Contraception and abortion do not guarantee reproductive autonomy, because many people cannot afford to have (and properly raise) any children, or as many children as they would like; and others are involuntarily infertile. However, both contraception and abortion are essential if women are to have the modest degree of reproductive autonomy which is possible in the world as it is presently constructed.

In the long run, access to abortion is essential for the health and survival not just of individual women and families, but also that of the larger social and biological systems on which all our lives depend. Given the inadequacy of present methods of contraception and the lack of universal access to contraception, the avoidance of rapid population growth generally requires some use of abortion. Unless population growth rates are reduced in those impoverished societies where they remain high, malnutrition and starvation will become even more widespread than at present. There might still be enough food to feed all the people of the world, if only it were more equitably distributed. However, this cannot remain true indefinitely. Soil erosion and climatic changes brought about by the destruction of forests and the burning of fossil fuels threaten to reduce the earth's capacity for food production — perhaps drastically — within the next generation.

Yet opponents of abortion deny that abortion is necessary for the avoidance of such undesirable consequences. Some pregnancies are the result of rape or involuntary incest, but most result from apparently voluntary sexual behaviour. Thus, anti-abortionists often claim that women who seek abortions are 'refusing to take responsibility for their own actions'. In their view, women ought to avoid heterosexual intercourse unless they are prepared to complete any resulting pregnancy. But is this demand a reasonable one?

Heterosexual intercourse is not *biologically* necessary for women's — or men's — individual survival or physical health. On the contrary, women who are celibate or homosexual are less vulnerable to cervical cancer, AIDS, and other sexually transmitted diseases. Nor is it obvious that sex is necessary for the psychological health of either women or men, although the contrary belief is widespread. It is, however, something that many women find intensely pleasurable — a fact which is morally significant on most consequentialist theories. Furthermore, it is part of the form of life which the majority of women everywhere appear to prefer. In some places, lesbian

women are creating alternative forms of life which may better serve their needs. But for most heterosexual women, the choice of permanent celibacy is very difficult. In much of the world, it is very difficult for single women to support themselves (let alone support a family); and sexual intercourse is usually one of the 'duties' of married women.

In short, permanent celibacy is not a reasonable option to impose upon most women. And since all women are potentially vulnerable to rape, even those who are homosexual or celibate may face unwanted pregnancies. Hence, until there is a fully reliable and safe form of contraception, available to all women, the consequentialist arguments for abortion will remain strong. But these arguments will not persuade those who reject consequentialist moral theories. If abortion is inherently wrong, as many believe, then it cannot be justified as a means of avoiding undesirable consequences. Thus, we must also consider whether women have a moral right to seek abortion.

### III ABORTION AND WOMEN'S RIGHTS

Not all moral philosophers believe that there are such things as moral rights. Thus, it is important to say a bit here about what moral rights are; in section (viii) I will say more about why they are important. . . . Rights are not mysterious entities that we discover in nature; they are not, in fact, entities at all. To say that people have a right to life is to say, roughly, that they should never deliberately be killed or deprived of the necessities of life, unless the only alternative is some much greater evil. Rights are not absolute, but neither are they to be overridden for just any apparently greater good. For instance, one may kill in self-defence when there is no other way to protect oneself from death or serious harm unjustly inflicted; but one may not kill another person merely because others may gain a great deal from the victim's death.

Basic moral rights are those which all persons have, in contrast to those rights which depend upon particular circumstances, e.g. promises or legal contracts. The basic moral rights of persons are usually held to include the rights to life, liberty, self-determination, and freedom from the infliction of bodily harm. The prohibition of abortion appears to infringe upon all of these basic rights. Women's lives are endangered in at least two ways. Where abortion is illegal, women often seek unsafe illegal abortions; the World Health Organization estimates that over 200,000 women die from this cause each year. Many others die from involuntary childbirth, when abortion is unavailable, or when they are pressured not to use it. Of course, voluntary childbirth also involves some risk of death; but in the absence of coercion, there is no violation of the woman's right to life.

The denial of abortion also infringes upon women's rights to liberty, self-determination, and physical integrity. To be forced to bear a child is not just

an ‘inconvenience’, as opponents of abortion often claim. To carry a pregnancy to term is an arduous and risky undertaking, even when voluntary. To be sure, many women enjoy (much of) their pregnancies; but for those who remain pregnant against their will the experience is apt to be thoroughly miserable. And involuntary pregnancy and childbirth are only the beginning of the hardships caused by the denial of abortion. The woman must either keep the child or surrender it for adoption. To keep the child may make it impossible to continue her chosen life work, or to meet her other family obligations. To surrender the child means that she must live with the unhappy knowledge that she has a daughter or son for whom she cannot care, often cannot even know to be alive and well. Studies of women who have surrendered infants for adoption show that, for most, the separation from their child is a great and lasting grief.

Even if we accept the view that fetuses have a right to life, it is difficult to justify the imposition of such hardships upon unwilling individuals for the sake of fetal lives. As Judith Thomson pointed out in her much-discussed 1971 article, ‘A defense of abortion’, there is no other case in which the law requires individuals (who have been convicted of no crime) to sacrifice liberty, self-determination, and bodily integrity in order to preserve the lives of others. Perhaps one analogy to involuntary childbirth is military conscription. However, that comparison can lend only moderate support to the anti-abortion position, since the justifiability of compulsory military service is itself debatable.

In popular rhetoric, especially in the United States, the abortion issue is often seen as purely and simply one of ‘women’s right to control their bodies’. If women have the moral right to abort unwanted pregnancies, then the law should not prohibit abortion. But the arguments for this right do not entirely solve the moral issue of abortion. For it is one thing to have a right, and another to be morally justified in exercising that right in a particular case. If fetuses have a full and equal right to life, then perhaps women’s right to abort should be exercised only in extreme circumstances. And perhaps we should question further whether fertile human beings — of either sex — are entitled to engage in heterosexual intercourse when they are not willing to have a child and assume the responsibility for it. If popular heterosexual activities are costing the lives of millions of innocent ‘persons’ (i.e. aborted fetuses), then should we not at least try to give up these activities? On the other hand, if fetuses do not yet have a substantial right to life, then abortion is not nearly so difficult to justify.

#### IV QUESTIONS ABOUT THE MORAL STATUS OF FETUSES

When in the development of a human individual does she or he begin to have a full and equal right to life? Most contemporary legal systems treat birth as the point at which a new legal person comes into existence. Thus, infanticide is generally classified as a form of homicide, whereas abortion — even where prohibited — generally is not. But, at first glance, birth seems to

be an entirely arbitrary criterion of moral status. Why should human beings attain full and equal basic moral rights at birth, rather than at some earlier or later point?

Many theorists have sought to establish some universal criterion of moral status, by which to distinguish between those entities that have full moral rights and those that have no moral rights, or different and lesser rights. Even those who prefer not to speak of moral rights may feel the need for a universally applicable criterion of moral status. For instance, utilitarians need to know which entities have interests that must be considered in calculations of moral utility, while Kantian deontologists need to know which things are to be treated as ends in themselves, and not merely means to the ends of others. Many criteria of moral status have been proposed. The most common include life, sentience (the capacity to have experiences, including that of pain), genetic humanity (biological identification as belonging to the species *Homo sapiens*), and personhood (which will be defined later).

How are we to choose among these conflicting criteria of moral status? Two things are clear. First, we may not treat the selection of a criterion of moral status as a mere matter of personal preference. Racists, for instance, are not entitled to recognize the moral rights only of members of their own racial group, since they have never been able to prove that members of 'inferior' races lack any property that can reasonably be held to be relevant to moral status. Second, a theory of moral status must provide a plausible account of the moral status not only of human beings, but also of non-human animals, plants, computers, possible extraterrestrial life forms, and anything else that might come along. I will argue that life, sentience and personhood are all relevant to moral status, though not in the same ways. Let us consider these criteria in turn, beginning with the most basic, i.e. biological life.

#### V THE ETHIC OF 'REVERENCE FOR LIFE'

Albert Schweitzer argued for an ethic of reverence for all living things. He held that all organisms, from microbes to human beings, have a 'will to live'. Thus, he says, anyone who has 'an unblunted moral sensibility will find it natural to share concern with the fate of all living creatures', Schweitzer may have been wrong to claim that all living things have a *will* to live. Will is most naturally construed as a faculty which requires at least some capacity for thought, and is thus unlikely to be found in simple organisms that lack central nervous systems. Perhaps the claim that all living things share a will to live is a metaphorical statement of the fact that organisms are teleologically organized, such that they generally function in ways that promote their own survival or that of their species. But why should this fact lead us to feel a reverence for all life?

I suggest that the ethic of reverence for life draws strength from ecological and aesthetic concerns. The destruction of living things often damages what

Aldo Leopold has called the ‘integrity, stability and beauty of the biotic community’. Protecting the biotic community from needless damage is a moral imperative, not just for the good of humanity, but because the unspoiled natural world is worth valuing for its own sake.

Reverence for life suggests that, other things being equal, it is always better to avoid killing a living thing. But Schweitzer was aware that not all killing can be avoided. His view was that one should never kill without good reason, and certainly not for sport or amusement. Thus, it does not follow from an ethic of reverence for all life that abortion is morally wrong. Human fetuses are living things, as are unfertilized ova and spermatazoa. However, many abortions may be defended as killing ‘under the compulsion of necessity’.

#### **VI GENETIC HUMANITY**

Opponents of abortion will reply that abortion is wrong, not simply because human fetuses are *alive*, but because they are *human*. But why should we believe that the destruction of a living human organism is always morally worse than the destruction of an organism of some other species? Membership in a particular biological species does not appear to be, in itself, any more relevant to moral status than membership in a particular race or sex.

It is an accident of evolution and history that everyone whom we currently recognize as having full and equal basic moral rights belongs to a single biological species. The ‘people’ of the earth might just as well have belonged to many different species — and indeed perhaps they do. It is quite possible that some non-human animals, such as dolphins and whales and the great apes, have enough so-called ‘human’ capacities to be properly regarded as persons — i.e. beings capable of reason, self-awareness, social involvement, and moral reciprocity. Some contemporary philosophers have argued that (some) non-human animals have essentially the same basic moral rights as human persons. Whether or not they are right, it is certainly true that any superior moral status accorded to members of our own species must be justified in terms of morally significant differences between humans and other living things. To hold that species alone provides a basis for superior moral status is arbitrary and unhelpful.

#### **VII THE SENTIENCE CRITERION**

Some philosophers hold that sentience is the primary criterion of moral status. Sentience is the capacity to have experiences — for instance, visual, auditory, olfactory, or other perceptual experiences. However, the capacity to have pleasurable and painful experiences seems particularly relevant to moral status. It is a plausible postulate of utilitarian ethics that pleasure is an intrinsic good and pain an intrinsic evil. True, the capacity to feel pain is often valuable to an organism, enabling it to avoid harm or destruction. Conversely, some pleasures can be harmful to the organism’s long-term

well-being. Nevertheless, sentient beings may be said to have a basic interest in pleasure and the avoidance of pain. Respect for this basic interest is central to utilitarian ethics.

The sentience criterion suggests that, other things being equal, it is morally worse to kill a sentient than a non-sentient organism. The death of a sentient being, even when painless, deprives it of whatever pleasurable experiences it might have enjoyed in the future. Thus, death is apt to be a misfortune for that being, in a way that the death of a non-sentient organism is not.

But how can we know which living organisms are sentient? For that matter, how can we know that non-living things, such as rocks and rivers, are not sentient? If knowledge requires the absolute impossibility of error, then we probably cannot know this. But what we do know strongly suggests that sentience requires a functioning central nervous system — which is absent in rocks, plants, and simple micro-organisms. It is also absent in the early human fetus. Many neuro-physiologists believe that normal human fetuses begin to have some rudimentary capacity for sentience at some stage in the second trimester of pregnancy. Prior to that stage, their brains and sensory organs are too undeveloped to permit the occurrence of sensations. The behavioural evidence points in the same direction. By the end of the first trimester, a fetus may have some unconscious reflexes, but it does not yet respond to its environment in a way suggestive of sentience. By the third trimester, however, some parts of the fetal brain are functional, and the fetus may respond to noise, light, pressure, motion and other sensory stimuli.

The sentience criterion lends support to the common belief that late abortion is more difficult to justify than early abortion. Unlike the presentient fetus, a third-trimester fetus is already a *being* — already, that is, a centre of experience. If killed, it may experience pain. Moreover, its death (like that of any sentient being) will mean the termination of a stream of experiences, some of which may have been pleasurable. Indeed, the use of this criterion suggests that early abortion poses no very serious moral issue, at least with respect to the impact upon the fetus. As a living but non-sentient organism, the first-trimester fetus is not yet a being with an interest in continued life. Like the unfertilized ovum, it may have the potential to *become* a sentient being. But this means only that it has the potential to become a being with an interest in continued life, not that it already has such an interest.

While the sentience criterion implies that late abortion is more difficult to justify than early abortion, it does not imply that late abortion is as difficult to justify as homicide. The principle of respect for the interests of sentient beings does not imply that all sentient beings have an *equal* right to life. To see why this is so, we need to give further thought to the scope of that principle.

Most normal mature vertebrate animals (mammals, birds, reptiles, amphibians and fish) are obviously sentient. It is also quite likely that many invertebrate animals, such as arthropods (e.g. insects, spiders, and crabs), are sentient. For they too have sense organs and nervous systems, and often behave as if they could see, hear, and feel quite well. If sentience is

the criterion of moral status, then not even a fly should be killed without some good reason.

But what counts as a good enough reason for the destruction of a living thing whose primary claim to moral status is its probable sentience? Utilitarians generally hold that acts are morally wrong if they increase the total amount of pain or suffering in the world (without some compensatory increase in the total amount of pleasure or happiness), or vice versa. But the killing of a sentient being does not always have such adverse consequences. There is room in any environment for only a finite number of organisms of any given species. When a rabbit is killed (in some more or less painless fashion), another rabbit is likely to take its place, so that the total amount of rabbit-happiness is not decreased. Moreover, rabbits, like many other rapidly reproducing species, must be preyed upon by some other species if the health of the larger biological system is to be maintained.

Thus, the killing of sentient beings is not always an evil in utilitarian terms. However, it would be morally offensive to suggest that people can be killed just because they are too numerous, and are upsetting the natural ecology. If killing people is harder to justify than killing rabbits — as even most animal liberationists believe — it must be because people have some moral status that is not based upon sentience alone. In the next section, we consider some possible arguments for this view.

#### VIII PERSONHOOD AND MORAL RIGHTS

Once they are past infancy, human beings typically possess not only a capacity for sentience, but also such ‘higher’ mental capacities as self-awareness and rationality. They are also highly social beings, capable — except in pathological cases — of love, nurturance, co-operation, and moral responsibility (which involves the capacity to guide their actions through moral principles and ideals). Perhaps these mental and social capacities can provide sound reasons for ascribing a stronger right to life to persons than to other sentient beings.

One argument for that conclusion is that the distinctive capacities persons have enable them to value their own lives and those of other members of their communities more than other animals do. People are the only beings who can plan years into the future, and who are often haunted by the fear of premature death. Perhaps this means that the lives of persons are worth more to their possessors than those of sentient non-persons. If so, then killing a person is a greater moral wrong than killing a sentient being which is not a person. But it is also possible that the absence of fear for the future tends to make the lives of sentient non-persons more pleasant, and more valuable to them, than ours are to us. Thus, we need to look elsewhere for a rationale for the superior moral status that most (human) persons accord to one another.

Moral rights are a way of talking about how we should behave. That it is evidently only persons who understand the idea of a moral right does not make us ‘better’ than other sentient beings. However, it does give us

compelling reasons for treating one another as moral equals, with basic rights that cannot be overridden for narrowly utilitarian reasons. If we could never trust other persons not to kill us whenever they judged that some net good might result, social relationships would become immeasurably more difficult, and the lives of all but the most powerful persons would be greatly impoverished.

A morally sensitive person will respect all life-forms, and will be careful to avoid needlessly inflicting pain or death upon sentient beings. However, she will respect the basic moral rights of other persons *as equal to her own*, not just because they are alive and sentient, but also because she can reasonably hope and demand that they will show her the same respect. Mice and mosquitoes are not capable of this kind of moral reciprocity — as least not in their interactions with human beings. When their interests come into conflict with ours, we cannot hope to use moral argument to persuade them to accept some reasonable compromise. Thus, it is often impossible to accord them fully equal moral status. Even the Jain religion of India, which regards the killing of any being as an obstacle to spiritual enlightenment, does not require the total avoidance of such killing, except in the case of those who have taken special religious vows. . . .

If the capacity for moral reciprocity is essential to personhood, and if personhood is the criterion for moral equality, then human fetuses do not satisfy that criterion. Sentient fetuses are closer to being persons than are fertilized ova or early fetuses, and may gain some moral status on that account. However, they are not yet reasoning, self-aware beings, capable of love, nurturance, and moral reciprocity. These facts lend support to the view that even late abortion is not quite the equivalent of homicide. On this basis, we may reasonably conclude that the abortion of sentient fetuses can sometimes be justified for reasons that could not justify the killing of a person. For instance, late abortion may sometimes be justified because the fetus has been found to be severely abnormal, or because the continuation of the pregnancy threatens the woman's health, or creates other personal hardships.

Unfortunately, the discussion cannot end at this point. Personhood is important as an *inclusion* criterion for moral equality: any theory which denies equal moral status to certain persons must be rejected. But personhood seems somewhat less plausible as an *exclusion* criterion, since it appears to exclude infants and mentally handicapped individuals who may lack the mental and social capacities typical of persons. Furthermore — as opponents of abortion point out — history proves that it is all too easy for dominant groups to rationalize oppression by claiming, in effect, that the oppressed persons are not really persons at all, because of some alleged mental or moral deficiency.

In view of these points it may seem wise to adopt the theory that all *sentient* human beings have full and equal basic moral rights. (To avoid 'speciesism', we could grant the same moral status to sentient members of any other species whose normal, mature members we believe to be persons.) On

this theory, so long as an individual is both human and capable of sentience, his or her moral equality cannot be questioned. But there is an objection to the extension of equal moral status even to *sentient* fetuses: it is impossible in practice to grant equal moral rights to fetuses without denying those same rights to women.

#### IX WHY BIRTH MATTERS MORALLY

There are many instances in which the moral rights of different human individuals come into apparent conflict. Such conflicts cannot, as a rule, be solved justly by denying equal moral status to one of the parties. But pregnancy is a special case. Because of the unique biological relationship between the woman and the fetus, the extension of equal moral and legal status to fetuses has ominous consequences for women's basic rights.

One consequence is that abortion 'on request' would not be permitted. If sentience is the criterion, then abortion might be permitted only in the first trimester. Some argue that this is a reasonable compromise, since it would allow most women time enough to discover that they are pregnant, and decide whether or not to abort. But problems involving fetal abnormality, the woman's health, or her personal or economic situation, sometimes arise or become severe only at a later stage. If fetuses are presumed to have the same moral rights as already-born human beings, then women will often be compelled to remain pregnant at great risk to their own lives, health, or personal well-being. They may also be compelled to submit, against their will, to dangerous and invasive medical procedures such as Caesarean section, when others judge that this would be beneficial to the fetus. (A number of such cases have already occurred in the United States.) Thus, the extension of full and equal basic moral rights to fetuses endangers the basic rights of women.

But, given these apparent conflicts between fetal rights and women's rights, one may still wonder why it is women's rights that should prevail. Why not favour fetuses instead, e.g. because they are more helpless, or have a longer life expectancy? Or why not seek a compromise between fetal and maternal rights, with equal concessions on each side? If fetuses were already persons, in the sense I have described, then it would be arbitrary to favour the rights of woman over theirs. But it is difficult to argue that either fetuses or newborn infants are persons in this sense, since the capacities for reason, self-awareness, and social and moral reciprocity seem to develop only after birth.

Why, then, should we treat birth, rather than some later point, as the threshold of moral equality? A major reason is that birth makes it possible for the infant to be granted equal basic rights without violating anyone else's basic rights. It is possible in many countries to find good homes for most infants whose biological parents are unable or unwilling to raise them. Since most of us strongly desire to protect infants, and since we can now do so without imposing excessive hardships upon women and families, there is no evident reason why we should not. But fetuses are different: their equality would mean women's inequality. Other things being equal, it is worse

to deny the basic moral rights to beings that are clearly not yet full persons. Since women are persons and fetuses are not, we should come down on the side of respecting women's rights in cases of apparent conflict.

#### X POTENTIAL PERSONHOOD

Some philosophers argue that, although fetuses may not be persons, their potential to *become* persons gives them the same basic moral rights. This argument is implausible, since in no other case do we treat the potential to achieve some status entailing certain rights as itself entailing those same rights. For instance, every child born in the United States is a potential voter, but no-one under the age of 18 has the right to vote in that country. Besides, the argument from potential proves too much. If a fetus is a potential person, then so is an unfertilized human ovum, together with enough viable spermatozoa to achieve fertilization; yet few would seriously suggest that *these* living human entities should have full and equal moral status.

Yet the argument from fetal potential refuses to go away. Perhaps this is because the potential which fetuses have is often a sound reason for valuing and protecting them. Once a pregnant woman has committed herself to the continued nurturance of the fetus, she and those close to her are likely to think of it as an 'unborn baby', and to value it for its potential. The fetus's potential lies not just in its DNA, but in that maternal (and paternal) commitment. Once the woman has committed herself to the pregnancy, it is appropriate for her to value the fetus and protect its potential — as most women do, without any legal coercion. But it is wrong to demand that a woman complete a pregnancy when she is unable or unwilling to undertake that enormous commitment.

#### XI SUMMARY AND CONCLUSION

Abortion is often approached as if it were *only* an issue of fetal rights; and often as if it were *only* an issue of women's rights. The denial of safe and legal abortion infringes upon women's rights to life, liberty, and physical integrity. Yet if the fetus had the same right to life as a person, abortion would still be a tragic event, and difficult to justify except in the most extreme cases. Thus, even those who argue for women's rights must be concerned with the moral status of fetuses.

Even an ethic of reverence for all life does not, however, preclude all intentional killing. All killing requires justification, and it is somewhat more difficult to justify the deliberate destruction of a sentient being than of a living thing which is not (yet) a centre of experience; but sentient beings do not all have equal rights. The extension of equal moral status to fetuses threatens women's most basic rights. Unlike fetuses, women are already persons. They should not be treated as something less when they happen to be pregnant. That is why abortion should not be prohibited, and why birth, rather than some earlier point, marks the beginning of full moral status.

**REFERENCES**

- Jaini, P.: *The Jaina Path of Purification* (Berkeley: University of California Press, 1979).
- Leopold, A.: *A Sand County Almanac* (New York: Ballantine Books, 1970).
- Schweitzer, A.: *The Teaching of Reverence for Life*, trans. R. and C. Winston (New York: Holt, Rinehart and Winston, 1965).
- Thomson, J. J.: 'A defense of abortion', *Philosophy and Public Affairs* 1:1 (Fall 1971), 47–66.

**DISCUSSION QUESTIONS**

1. What are the main consequentialist arguments for abortion? Why does Warren think these arguments are inadequate? Do you agree? Why or why not?
2. Warren writes, "it is one thing to have a right, and another to be morally justified in exercising that right in a particular case." Which rights and cases is she talking about? Why does she draw this distinction?
3. Does Warren argue for a single criterion of moral status? If so, what is it? If not, is this a problem for her argument?
4. Is it "other things being equal, . . . always better to avoid killing a living thing"? Why or why not? What does this claim imply about abortion?
5. According to Warren, "To hold that species alone provides a basis for superior moral status is arbitrary and unhelpful." Why does she think this? Do you agree? Why or why not?
6. Is late abortion more difficult to justify than early abortion? Why or why not?
7. What does Warren mean by "the capacity for moral reciprocity"? Is this capacity sufficient for personhood? Is it necessary for personhood? Why or why not?
8. Why does birth matter morally, according to Warren? Do you agree? Why or why not?
9. Explain what Warren means when she says, "The fetus's potential lies not just in its DNA, but in that maternal and paternal commitment." What does she mean by "potential" here? Do opponents of abortion mean the same thing by "potential" when they use this concept to argue against abortion?
10. In another famous article, "On the Moral and Legal Status of Abortion" (*The Monist* 57, no. 1 (1973): 43–61), Warren imagines the following example:  
Suppose that our space explorer falls into the hands of an alien culture, whose scientists decide to create a few hundred thousand or more human beings, by breaking his body into its component cells, and using these to create fully developed human beings with, of course, his genetic code.

We may imagine that each of these newly created men will . . . be a bona fide (though hardly unique) person. Imagine that the whole project will take only seconds, and that its chances of success are extremely high, and that our explorer knows all of this, and also knows that these people will be treated fairly. I maintain that in such a situation he would have every right to escape if he could, and thus to deprive all of these potential people of their potential lives. . . ."

Warren cites this example in order to show that "the rights of any actual person invariably outweigh the rights of any [merely] potential person." Reconstruct (or construct!) her argument from her example to her conclusion. What, if anything, does this claim about rights show about abortion? How could opponents best respond to this argument?

11. Do Warren's conclusions about *morality* support any further conclusions about *laws* restricting abortions?

## AN ARGUMENT THAT ABORTION IS WRONG<sup>5</sup>

by Don Marquis

The purpose of this essay is to set out an argument for the claim that abortion, except perhaps in rare instances, is seriously wrong. One reason for these exceptions is to eliminate from consideration cases whose ethical analysis should be controversial and detailed for clear-headed opponents of abortion. Such cases include abortion after rape and abortion during the first fourteen days after conception when there is an argument that the fetus is not definitely an individual. Another reason for making these exceptions is to allow for those cases in which the permissibility of abortion is compatible with the argument of this essay. Such cases include abortion when continuation of a pregnancy endangers a woman's life and abortion when the fetus is anencephalic. When I speak of the wrongness of abortion in this essay, a reader should presume the above qualifications. I mean by an abortion an action intended to bring about the death of a fetus for the sake of the woman who carries it. (Thus, as is standard on the literature on this subject, I eliminate spontaneous abortions from consideration.) I mean by a fetus a developing human being from the time of conception to the time of birth. (Thus, as is standard, I call embryos and zygotes, fetuses.)

The argument of this essay will establish that abortion is wrong for the same reason as killing a reader of this essay is wrong. I shall just assume, rather than establish, that killing you is seriously wrong. I shall make no attempt to offer a complete ethics of killing. Finally, I shall make no attempt to resolve some very fundamental and difficult general philosophical issues into which this analysis of the ethics of abortion might lead. . . .

Source: Don Marquis, "Why Abortion is Immoral," *The Journal of Philosophy*, vol. 86, no. 4 (April 1989), pp. 183–202. Reprinted by permission.

**THE “FUTURE LIKE OURS” ACCOUNT OF THE WRONGNESS OF KILLING**

... Opponents of abortion claim that abortion is wrong because abortion involves killing someone like us, a human being who just happens to be very young. Supporters of choice claim that ending the life of a fetus is not in the same moral category as ending the life of an adult human being. Surely this controversy cannot be resolved in the absence of an account of what it is about killing us that makes killing us wrong. On the one hand, if we know what property we possess that makes killing us wrong, then we can ask whether fetuses have the same property. On the other hand, suppose that we do not know what it is about us that makes killing us wrong. If this is so, we do not understand even easy cases in which killing is wrong. Surely, we will not understand the ethics of killing fetuses, for if we do not understand easy cases, then we will not understand hard cases. Both pro-choicer and anti-abortionist agree that it is obvious that it is wrong to kill us. Thus, a discussion of what it is about us that makes killing us not only wrong, but seriously wrong, seems to be the right place to begin a discussion of the abortion issue.

Who is primarily wronged by a killing? The wrong of killing is not primarily explained in terms of the loss to the family and friends of the victim. Perhaps the victim is a hermit. Perhaps one's friends find it easy to make new friends. The wrong of killing is not primarily explained in terms of the brutalization of the killer. The great wrong to the victim explains the brutalization, not the other way around. The wrongness of killing us is understood in terms of what killing does to us. Killing us imposes on us the misfortune of premature death. That misfortune underlies the wrongness.

Premature death is a misfortune because when one is dead, one has been deprived of life. This misfortune can be more precisely specified. Premature death cannot deprive me of my past life. That part of my life is already gone. If I die tomorrow or if I live thirty more years my past life will be no different. It has occurred on either alternative. Rather than my past, my death deprives me of my future, of the life that I would have lived if I had lived out my natural life span.

The loss of a future biological life does not explain the misfortune of death. Compare two scenarios: in the former I now fall into a coma from which I do not recover until my death in thirty years. In the latter I die now. The latter scenario does not seem to describe a greater misfortune than the former.

The loss of our future conscious life is what underlies the misfortune of premature death. Not any future conscious life qualifies, however. Suppose that I am terminally ill with cancer. Suppose also that pain and suffering would dominate my future conscious life. If so, then death would not be a misfortune for me.

Thus, the misfortune of premature death consists of the loss to us of the future goods of consciousness. What are these goods? Much can be said about this issue, but a simple answer will do for the purposes of this essay. The goods of life are whatever we get out of life. The goods of life are those

items toward which we take a “pro” attitude. They are completed projects of which we are proud, the pursuit of our goals, aesthetic enjoyments, friendships, intellectual pursuits, and physical pleasures of various sorts. The goods of life are what makes life worth living. In general, what makes life worth living for one person will not be the same as what makes life worth living for another. Nevertheless, the list of goods in each of our lives will overlap. The lists are usually different in different stages of our lives.

What makes the goods of my future good for me? One possible, but wrong, answer is my desire for those goods now. This answer does not account for those aspects of my future life that I now believe I will later value, but about which I am wrong. Neither does it account for those aspects of my future that I will come to value, but which I don’t value now. What is valuable to the young may not be valuable to the middle-aged. What is valuable to the middle-aged may not be valuable to the old. Some of life’s values for the elderly are best appreciated by the elderly. Thus it is wrong to say that the value of my future to me is just what I value now. What makes my future valuable to me are those aspects of my future that I will (or would) value when I will (or would) experience them, whether I value them now or not.

It follows that a person can believe that she will have a valuable future and be wrong. Furthermore, a person can believe that he will not have a valuable future and also be wrong. This is confirmed by our attitude toward many of the suicidal. We attempt to save the lives of the suicidal and to convince them that they have made an error in judgment. This does not mean that the future of an individual obtains value from the value that others confer on it. It means that, in some cases, others can make a clearer judgment of the value of a person’s future *to that person* than the person herself. This often happens when one’s judgment concerning the value of one’s own future is clouded by personal tragedy. (Compare the views of McInerney, 1990, and Shirley, 1995.)

Thus, what is sufficient to make killing us wrong, in general, is that it causes premature death. Premature death is a misfortune. Premature death is a misfortune, in general, because it deprives an individual of a future of value. An individual’s future will be valuable to that individual if that individual will come, or would come, to value it. We know that killing us is wrong. What makes killing us wrong, in general, is that it deprives us of a future of value. Thus, killing someone is wrong, in general, when it deprives her of a future like ours. I shall call this “an FLO.”

#### ARGUMENTS IN FAVOR OF THE FLO THEORY

At least four arguments support this FLO account of the wrongness of killing.

##### *The Considered Judgment Argument*

The FLO account of the wrongness of killing is correct because it fits with our considered judgment concerning the nature of the misfortune of death.

The analysis of the previous section is an exposition of the nature of this considered judgment. This judgment can be confirmed. If one were to ask individuals with AIDS or with incurable cancer about the nature of their misfortune, I believe that they would say or imply that their impending loss of an FLO makes their premature death a misfortune. If they would not, then the FLO account would plainly be wrong.

#### *The Worst of Crimes Argument*

The FLO account of the wrongness of killing is correct because it explains why we believe that killing is one of the worst of crimes. My being killed deprives me of more than does my being robbed or beaten or harmed in some other way because my being killed deprives me of all of the value of my future, not merely part of it. This explains why we make the penalty for murder greater than the penalty for other crimes.

As a corollary the FLO account of the wrongness of killing also explains why killing an adult human being is justified only in the most extreme circumstances, only in circumstances in which the loss of life to an individual is outweighed by a worse outcome if that life is not taken. Thus, we are willing to justify killing in self-defense, killing in order to save one's own life, because one's loss if one does not kill in that situation is so very great. We justify killing in a just war for similar reasons. We believe that capital punishment would be justified if, by having such an institution, fewer premature deaths would occur. The FLO account of the wrongness of killing does not entail that killing is always wrong. Nevertheless, the FLO account explains both why killing is one of the worst of crimes and, as a corollary, why the exceptions to the wrongness of killing are so very rare. A correct theory of the wrongness of killing should have these features.

#### *The Appeal to Cases Argument*

The FLO account of the wrongness of killing is correct because it yields the correct answers in many life-and-death cases that arise in medicine and have interested philosophers.

Consider medicine first. Most people believe that it is not wrong deliberately to end the life of a person who is permanently unconscious. Thus we believe that it is not wrong to remove a feeding tube or a ventilator from a permanently comatose patient, knowing that such a removal will cause death. The FLO account of the wrongness of killing explains why this is so. A patient who is permanently unconscious cannot have a future that she would come to value, whatever her values. Therefore, according to the FLO theory of the wrongness of killing, death could not, *ceteris paribus*, be a misfortune to her. Therefore, removing the feeding tube or ventilator does not wrong her.

By contrast, almost all people believe that it is wrong, *ceteris paribus*, to withdraw medical treatment from patients who are temporarily unconscious. The FLO account of the wrongness of killing also explains why this

is so. Furthermore, these two unconsciousness cases explain why the FLO account of the wrongness of killing does not include present consciousness as a necessary condition for the wrongness of killing.

Consider now the issue of the morality of legalizing active euthanasia. Proponents of active euthanasia argue that if a patient faces a future of intractable pain and wants to die, then, *ceteris paribus*, it would not be wrong for a physician to give him medicine that she knows would result in his death. This view is so universally accepted that even the strongest *opponents* of active euthanasia hold it. The official Vatican view (Sacred Congregation, 1980) is that it is permissible for a physician to administer to a patient morphine sufficient (although no more than sufficient) to control his pain even if she foresees that the morphine will result in his death. Notice how nicely the FLO account of the wrongness of killing explains this unanimity of opinion. A patient known to be in severe intractable pain is presumed to have a future without positive value. Accordingly, death would not be a misfortune for him and an action that would (foreseeably) end his life would not be wrong.

Contrast this with the standard emergency medical treatment of the suicidal. Even though the suicidal have indicated that they want to die, medical personnel will act to save their lives. This supports the view that it is not the mere *desire* to enjoy an FLO which is crucial to our understanding of the wrongness of killing. *Having* an FLO is what is crucial to the account, although one would, of course, want to make an exception in the case of fully autonomous people who refuse life-saving medical treatment. Opponents of abortion can, of course, be willing to make an exception for fully autonomous fetuses who refuse life support.

The FLO theory of the wrongness of killing also deals correctly with issues that have concerned philosophers. It implies that it would be wrong to kill (peaceful) persons from outer space who come to visit our planet even though they are biologically utterly unlike us. Presumably, if they are persons, then they will have futures that are sufficiently like ours so that it would be wrong to kill them. The FLO account of the wrongness of killing shares this feature with the personhood views of the supporters of choice. Classical opponents of abortion who locate the wrongness of abortion somehow in the biological humanity of a fetus cannot explain this.

The FLO account does not entail that there is another species of animals whose members ought not to be killed. Neither does it entail that it is permissible to kill any non-human animal. On the one hand, a supporter of animals' rights might argue that since some non-human animals have a future of value, it is wrong to kill them also, or at least it is wrong to kill them without a far better reason than we usually have for killing non-human animals. On the other hand, one might argue that the futures of non-human animals are not sufficiently like ours for the FLO account to entail that it is wrong to kill them. Since the FLO account does not specify which properties a future of another individual must possess so that killing that individual is wrong, the FLO account is indeterminate with respect to this issue. The fact that the

FLO account of the wrongness of killing does not give a determinate answer to this question is not a flaw in the theory. A sound ethical account should yield the right answers in the obvious cases; it should not be required to resolve every disputed question.

A major respect in which the FLO account is superior to accounts that appeal to the concept of person is the explanation the FLO account provides of the wrongness of killing infants. There was a class of infants who had futures that included a class of events that were identical to the futures of the readers of this essay. Thus, reader, the FLO account explains why it was as wrong to kill you when you were an infant as it is to kill you now. This account can be generalized to almost all infants. Notice that the wrongness of killing infants can be explained in the absence of an account of what makes the future of an individual sufficiently valuable so that it is wrong to kill that individual. The absence of such an account explains why the FLO account is indeterminate with respect to the wrongness of killing non-human animals.

If the FLO account is the correct theory of the wrongness of killing, then because abortion involves killing fetuses and fetuses have FLOs for exactly the same reasons that infants have FLOs, abortion is presumptively seriously immoral. This inference lays the necessary groundwork for a fourth argument in favor of the FLO account that shows that abortion is wrong.

#### *The Analogy with Animals Argument*

Why do we believe it is wrong to cause animals suffering? We believe that, in our own case and in the case of other adults and children, suffering is a misfortune. It would be as morally arbitrary to refuse to acknowledge that animal suffering is wrong as it would be to refuse to acknowledge that the suffering of persons of another race is wrong. It is, on reflection, suffering that is a misfortune, not the suffering of white males or the suffering of humans. Therefore, infliction of suffering is presumptively wrong no matter on whom it is inflicted and whether it is inflicted on persons or nonpersons. Arbitrary restrictions on the wrongness of suffering count as racism or speciesism. Not only is this argument convincing on its own, but it is the only way of justifying the wrongness of animal cruelty. Cruelty toward animals is clearly wrong. (This famous argument is due to Singer, 1979.)

The FLO account of the wrongness of abortion is analogous. We believe that, in our own case and the cases of other adults and children, the loss of a future of value is a misfortune. It would be as morally arbitrary to refuse to acknowledge that the loss of a future of value to a fetus is wrong as to refuse to acknowledge that the loss of a future of value to Jews (to take a relevant twentieth-century example) is wrong. It is, on reflection, the loss of a future of value that is a misfortune; not the loss of a future of value to adults or loss of a future of value to non-Jews. To deprive someone of a future of value is wrong no matter on whom the deprivation is inflicted and no matter whether the deprivation is inflicted on persons or nonpersons.

Arbitrary restrictions on the wrongness of this deprivation count as racism, genocide or ageism. Therefore, abortion is wrong. This argument that abortion is wrong should be convincing because it has the same form as the argument for the claim that causing pain and suffering to non-human animals is wrong. Since the latter argument is convincing, the former argument should be also. Thus, an analogy with animals supports the thesis that abortion is wrong.

#### REPLIES TO OBJECTIONS

The four arguments in the previous section establish that abortion is, except in rare cases, seriously immoral. Not surprisingly, there are objections to this view. There are replies to the four most important objections to the FLO argument for the immorality of abortion.

##### *The Potentiality Objection*

The FLO account of the wrongness of abortion is a potentiality argument. To claim that a fetus *has* an FLO is to claim that a fetus now has the potential to be in a state of a certain kind in the future. It is not to claim that all ordinary fetuses *will* have FLOs. Fetuses who are aborted, of course, will not. To say that a standard fetus has an FLO is to say that a standard fetus either will have or would have a life it will or would value. To say that a standard fetus would have a life it would value is to say that it will have a life it will value if it does not die prematurely. The truth of this conditional is based upon the nature of fetuses (including the fact that they naturally age) and this nature concerns their potential.

Some appeals to potentiality in the abortion debate rest on unsound inferences. For example, one may try to generate an argument against abortion by arguing that because persons have the right to life, potential persons also have the right to life. Such an argument is plainly invalid as it stands. The premise one needs to add to make it valid would have to be something like: "If Xs have the right to Y, then potential Xs have the right to Y." This premise is plainly false. Potential presidents don't have the rights of the presidency; potential voters don't have the right to vote.

In the FLO argument potentiality is not used in order to bridge the gap between adults and fetuses as is done in the argument in the above paragraph. The FLO theory of the wrongness of killing adults is based upon the adult's potentiality to have a future of value. Potentiality is in the argument from the very beginning. Thus, the plainly false premise is not required. Accordingly, the use of potentiality in the FLO theory is not a sign of an illegitimate inference.

##### *The Argument from Interests*

A second objection to the FLO account of the immorality of abortion involves arguing that even though fetuses have FLOs, nonsentient fetuses do not meet

the minimum conditions for having any moral standing at all because they lack interests. Steinbock (1992, p. 5) has presented this argument clearly:

Beings that have moral status must be capable of caring about what is done to them. They must be capable of being made, if only in a rudimentary sense, happy or miserable, comfortable or distressed. Whatever reasons we may have for preserving or protecting non-sentient beings, these reasons do not refer to their own interests. For without conscious awareness, beings cannot have interests. Without interests, they cannot have a welfare of their own. Without a welfare of their own, nothing can be done for their sake. Hence, they lack moral standing or status.

Medical researchers have argued that fetuses do not become sentient until after 22 weeks of gestation (Steinbock, 1992, p. 50). If they are correct, and if Steinbock's argument is sound, then we have both an objection to the FLO account of the wrongness of abortion and a basis for a view on abortion minimally acceptable to most supporters of choice.

Steinbock's conclusion conflicts with our settled moral beliefs. Temporarily unconscious human beings are nonsentient, yet no one believes that they lack either interests or moral standing. Accordingly, neither conscious awareness nor the capacity for conscious awareness is a necessary condition for having interests.

The counter-example of the temporarily unconscious human being shows that there is something internally wrong with Steinbock's argument. The difficulty stems from an ambiguity. One cannot *take* an interest in something without being capable of caring about what is done to it. However, something can be *in* someone's interest without that individual being capable of caring about it, or about anything. Thus, life support can be *in* the interests of a temporarily unconscious patient even though the temporarily unconscious patient is incapable of *taking* an interest in that life support. If this can be so for the temporarily unconscious patient, then it is hard to see why it cannot be so for the temporarily unconscious (that is, non-sentient) fetus who requires placental life support. Thus the objection based on interests fails.

### *The Problem of Equality*

The FLO account of the wrongness of killing seems to imply that the degree of wrongness associated with each killing varies inversely with the victim's age. Thus, the FLO account of the wrongness of killing seems to suggest that it is far worse to kill a five-year-old than an 89-year-old because the former is deprived of far more than the latter. However, we believe that all persons have an equal right to life. Thus, it appears that the FLO account of the wrongness of killing entails an obviously false view (Paske, 1994).

However, the FLO account of the wrongness of killing does not, strictly speaking, imply that it is worse to kill younger people than older people. The FLO account provides an explanation of the wrongness of killing that

is sufficient to account for the serious presumptive wrongness of killing. It does not follow that killings cannot be wrong in other ways. For example, one might hold, as does Feldman (1992, p. 184), that in addition to the wrongness of killing that has its basis in the future life of which the victim is deprived, killing an individual is also made wrong by the admirability of an individual's past behavior. Now the amount of admirability will presumably vary directly with age, whereas the amount of deprivation will vary inversely with age. This tends to equalize the wrongness of murder.

However, even if, *ceteris paribus*, it is worse to kill younger persons than older persons, there are good reasons for adopting a doctrine of the legal equality of murder. Suppose that we tried to estimate the seriousness of a crime of murder by appraising the value of the FLO of which the victim had been deprived. How would one go about doing this? In the first place, one would be confronted by the old problem of interpersonal comparisons of utility. In the second place, estimation of the value of a future would involve putting oneself, not into the shoes of the victim at the time she was killed, but rather into the shoes the victim would have worn had the victim survived, and then estimating from that perspective the worth of that person's future. This task seems difficult, if not impossible. Accordingly, there are reasons to adopt a convention that murders are equally wrong.

Furthermore, the FLO theory, in a way, explains why we do adopt the doctrine of the legal equality of murder. The FLO theory explains why we regard murder as one of the worst of crimes, since depriving someone of a future like ours deprives her of more than depriving her of anything else. This gives us a reason for making the punishment for murder very harsh, as harsh as is compatible with civilized society. One should not make the punishment for younger victims harsher than that. Thus, the doctrine of the equal legal right to life does not seem to be incompatible with the FLO theory.

### *The Contraception Objection*

The strongest objection to the FLO argument for the immorality of abortion is based on the claim that, because contraception results in one less FLO, the FLO argument entails that contraception, indeed, abstention from sex when conception is possible, is immoral. Because neither contraception nor abstention from sex when conception is possible is immoral, the FLO account is flawed.

There is a cogent reply to this objection. If the argument of the early part of this essay is correct, then the central issue concerning the morality of abortion is the problem of whether fetuses are individuals who are members of the class of individuals whom it is seriously presumptively wrong to kill. The properties of being human and alive, of being a person, and of having an FLO are criteria that participants in the abortion debate have offered to mark off the relevant class of individuals. The central claim of this essay is that having an FLO marks off the relevant class of individuals. A defender of the FLO view could, therefore, reply that since, at the time of contraception,

there is no individual to have an FLO, the FLO account does not entail that contraception is wrong. The wrong of killing is primarily a wrong to the individual who is killed; at the time of contraception there is no individual to be wronged.

However, someone who presses the contraception objection might have an answer to this reply. She might say that the sperm and egg are the individuals deprived of an FLO at the time of contraception. Thus, there are individuals whom contraception deprives of an FLO and if depriving an individual of an FLO is what makes killing wrong, then the FLO theory entails that contraception is wrong.

There is also a reply to this move. In the case of abortion, an objectively determinate individual is the subject of harm caused by the loss of an FLO. This individual is a fetus. In the case of contraception, there are far more candidates (see Norcross, 1990). Let us consider some possible candidates in order of the increasing number of individuals harmed: (1) The single harmed individual might be the combination of the particular sperm and the particular egg that would have united to form a zygote if contraception had not been used. (2) The two harmed individuals might be the particular sperm itself, and, in addition, the ovum itself that would have physically combined to form the zygote. (This is modeled on the double homicide of two persons who would otherwise in a short time fuse. (1) is modeled on harm to a single entity some of whose parts are not physically contiguous, such as a university.) (3) The many harmed individuals might be the millions of *combinations* of sperm and the released ovum whose (small) chances of having an FLO were reduced by the successful contraception. (4) The even larger class of harmed individuals (larger by one) might be the class consisting of all of the individual sperm in an ejaculate and, in addition, the individual ovum released at the time of the successful contraception. (1) through (4) are all candidates for being the subject(s) of harm in the case of successful contraception or abstinence from sex. Which should be chosen? Should we hold a lottery? There seems to be no non-arbitrarily determinate subject of harm in the case of successful contraception. But if there is no such subject of harm, then no determinate thing was harmed. If no determinate thing was harmed, then (in the case of contraception) no wrong has been done. Thus, the FLO account of the wrongness of abortion does not entail that contraception is wrong.

#### CONCLUSION

This essay contains an argument for the view that, except in unusual circumstances, abortion is seriously wrong. Deprivation of an FLO explains why killing adults and children is wrong. Abortion deprives fetuses of FLOs. Therefore, abortion is wrong. This argument is based on an account of the wrongness of killing that is a result of our considered judgment of the nature of the misfortune of premature death. It accounts for why we regard killing as one of the worst of crimes. It is superior to alternative

accounts of the wrongness of killing that are intended to provide insight into the ethics of abortion. This account of the wrongness of killing is supported by the way it handles cases in which our moral judgments are settled. This account has an analogue in the most plausible account of the wrongness of causing animals to suffer. This account makes no appeal to religion. Therefore, the FLO account shows that abortion, except in rare instances, is seriously wrong.

#### REFERENCES

- Feldman, F., *Confrontations with the Reaper. A Philosophical Study of the Nature and Value of Death* (New York: Oxford University Press, 1992).
- Marquis, D. B., "Why Abortion Is Immoral," *Journal of Philosophy* 86 (1989): 183–202.
- , "A Future Like Ours and the Concept of Person: A Reply to McInerney and Paske," *The Abortion Controversy: A Reader*, ed. L. P. Pojman and F. J. Beckwith (Boston: Jones and Bartlett, 1994), 354–68.
- , "Fetuses, Futures and Values: A Reply to Shirley," *Southwest Philosophy Review* 11 (1995): 263–65.
- McInerney, P., "Does a Fetus Already Have a Future Like Ours?" *Journal of Philosophy* 87 (1990): 264–68.
- Norcross, A., "Killing, Abortion, and Contraception: A Reply to Marquis," *Journal of Philosophy* 87 (1990): 268–77.
- Paske, G., "Abortion and the Neo-natal Right to Life: A Critique of Marquis's Futurist Argument," *The Abortion Controversy: A Reader*, ed. L. P. Pojman and F. J. Beckwith (Boston: Jones and Bartlett, 1994), 343–53.
- Sacred Congregation for the Propagation of the Faith, *Declaration on Euthanasia* (Vatican City, 1980).
- Shirley, E. S., "Marquis' Argument Against Abortion: A Critique," *Southwest Philosophy Review* 11 (1995): 79–89.
- Singer, P., "Not for Humans Only: The Place of Nonhumans in Environmental Issues," *Ethics and Problems of the 21st Century*, ed. K. E. Goodpaster and K. M. Sayre (South Bend: Notre Dame University Press, 1979).
- Steinbock, B., *Life Before Birth: The Moral and Legal Status of Embryos and Fetuses* (New York: Oxford University Press, 1992).

#### DISCUSSION QUESTIONS

1. Marquis's argument against abortion can be seen as an inference to the best explanation of why it is morally wrong to kill normal adult human beings. What does Marquis take to be the best explanation of this? Why does he think that it is better than the alternatives? Can you give an even better explanation that accounts for the cases that Marquis discusses?
2. Is Marquis's argument different in important ways from traditional arguments based on the premise that a fetus is a person or a potential person? If so, how? If not, why not?

3. What exactly does Marquis mean by his claim that a fetus has a future like ours? What is it to "have" a future? Do I already now have my future illnesses and successes? When is a future "like ours"? Which similarities matter? Why?
4. Explain the contraception objection to Marquis's argument and his response. Is his response adequate? Why or why not?
5. Determine whether you think abortion is morally wrong in the following cases:
  - a. Where the mother is in danger of dying if she does not have an abortion.
  - b. Where the pregnancy is due to rape.
  - c. Where contraception was used, but it failed.
  - d. Where the fetus has a disease that usually causes death within a year or two.
  - e. Where the fetus has a disease that usually causes severe mental retardation.
  - f. Where the pregnant woman is mentally or physically unable to be a good mother.
  - g. Where the mother will suffer severe personal losses if the pregnancy continues.

Now try to formulate principles and analogies to justify your positions in these controversial cases.

6. Which underlying principles, if any, could protect human lives with only a few exceptions, yet allow us to take lives of:
  - a. Contract killers sentenced to capital punishment.
  - b. Humans who are in irreversible comas.
  - c. Ourselves in suicide.
  - d. Children who live next to munitions factories that are bombed in a war.
  - e. Animals for food, clothing, and entertainment.
7. Describe a moral problem that you have faced in your personal life, and apply the methods of moral reasoning that you have learned in this chapter.

## NOTES

<sup>1</sup> For a discussion of this approach, and for an excellent model of how to argue about a moral principle, see Philippa Foot, "The Problem of Abortion and the Doctrine of Double Effect," in Philippa Foot, *Virtues and Vices and Other Essays in Moral Philosophy* (Berkeley: University of California Press, 1978), 19–32.

<sup>2</sup> Gradualism is discussed in more detail in Joel Feinberg, "Abortion," in *Matters of Life and Death*, 2nd ed., ed. Tom Regan (New York: Random House, 1986), 256–93; and Margaret Olivia Little, "Abortion and the Margins of Personhood," *Rutgers Law Journal* 39 (2008): 331–48.

<sup>3</sup> Judith Jarvis Thomson, "A Defense of Abortion," *Philosophy and Public Affairs* 1, no. 1 (Fall 1971): 47–66.

<sup>4</sup> From *A Companion to Ethics*, ed. Peter Singer (Oxford: Basil Blackwell, 1991), 303–16.

<sup>5</sup> From *Ethics in Practice: An Anthology*, 2nd ed., ed. Hugh LaFollette (Malden, MA: Blackwell, 2002), 83–93. This essay was originally published in the first edition of this anthology. This is an updated version of a view that first appeared in the *Journal of Philosophy* (1989). This essay incorporates attempts to deal with the objections of McInerney (1990), Norcross (1990), Shirley (1995), Steinbock (1992), and Paske (1994) to the original version of the view.



## SCIENTIFIC REASONING

*The products of science are all around us. We depend on science when we drive cars, listen to compact discs, and cook food in microwaves. Still, few people understand how science operates. To some, the scientific enterprise seems to consist of nothing more than amassing huge quantities of data to prove or disprove some hypothesis. There is more to science than that. Scientists also seek theories that are profound and far-reaching, or even elegant and beautiful, as is shown by how scientists themselves speak about their theories. The goal of scientific theory is not just to list and describe natural phenomena but to make sense of nature—that is, to explain it, make it more intelligible. To choose among conflicting scientific theories, we have to decide which theory makes the most sense and provides the best explanations. This chapter will bring out this complex nature of science by discussing the scientific enterprise in general and then focusing on a particular debate about biological evolution.*

### STANDARD SCIENCE

---

The beginning of science lies in observation. When we look at the world around us, we see that many things happen. Apples fall off trees, the leaves of some trees change color in the autumn, the tides come in and go out, chickens lay eggs, and so on. One job of scientists is to describe and classify what happens and what exists. But scientists also wonder *why* some things happen rather than others. Maple trees change color in the fall, and spruce trees do not, but why? Chickens lay eggs, and monkeys do not, but why? A sphere of wood floats in water, and a gold sphere does not, but why? And why does gold float when pressed into the shape of a boat? These questions ask for *explanations*.

To provide an explanation, scientists often give arguments of the kind discussed in Chapter 1. The event to be explained is derived from a general principle plus a statement of initial conditions or particular facts. For example, given the general principle that a solid sphere floats in water if and only if it is less dense than water, and also given the particular facts that wood is less dense than water, whereas gold is denser than water, we can explain why a wooden sphere floats in water and a gold sphere does not.

Scientists often seek *deeper* explanations by asking why certain general principles themselves are true. The principle that a sphere floats in water only when it is less dense than water can be explained as an instance of the more general principle that *anything* floats only when it displaces more than its own weight in water. This broader principle not only explains why a wooden sphere floats in water but also why a piece of gold will float when molded into the form of a boat. This broader principle is in turn explained by deriving it from even more basic principles about gravity and the mutual repulsion of molecules. A larger scientific theory is thus used to explain not only why particular things happen but also why certain general principles hold.

Of course, scientists often put forward conflicting theories, so we need some way to test which theory is correct. One simple method is to use the theory to make *predictions*. Since an explanation depends on principles that are general, these principles have implications beyond the particular phenomenon that they were originally intended to explain. The theory thus predicts what will happen in circumstances that the scientist has not yet observed. We can then test the theory by seeing whether these predictions hold true. For example, we can make spheres out of a wide variety of materials, calculate their densities, and then see which ones float. If any sphere that is denser than water floats, then we have to give up our principle that a sphere floats in water *only if* it is less dense than water. (This is an application of the necessary condition test, discussed in Chapter 10.) If we find a sphere that is less dense than water but does not float, then we have to give up the principle that a sphere floats if it is less dense than water. (This is an application of the sufficient condition test, discussed in Chapter 10.)

These methods help us rule out certain scientific principles, but the fact that a principle implies true predictions does not, by itself, prove that the principle is true. That argument would commit something like the fallacy of affirming the consequent (see Exercise XXI in Chapter 6). Nonetheless, we can still say that a theory is *confirmed* if it yields true predictions, and it is confirmed more strongly if it yields more, more varied, and more unexpected true predictions.

Scientific method is actually much more complex than this simple example suggests. This becomes apparent when we encounter *anomalies*. Suppose we have confirmed and explained the principle that a sphere floats in water if and only if it is less dense than water. Suppose also that another principle is well confirmed: A substance gets smaller and denser as it gets colder. Taken together, these principles predict that a sphere of ice should sink in water. Ice is colder than water, so, according to the second principle, ice should be denser than water, and that, given the first principle, means that it should not float in water. Of course, our prediction is wrong, since spheres of ice do float in water. What do we do now? The obvious solution is to modify the principle that a substance gets smaller and denser as it gets colder. This holds for most substances, but not for water. Water expands and thus gets less dense as it freezes.

We could have tried another solution. We could have denied the other principle: that a sphere floats in water if and only if it is less dense than water. Why do scientists not do this? One reason is that we have independent evidence that water expands when it freezes. That is why jars of water burst when they are left in a freezer. Another reason is that we could not give up this principle alone, because it follows from more basic principles about gravity and the mutual repulsion of molecules. Thus, many other areas of science would be affected if we gave up the principle that a sphere floats in water if and only if it is less dense than water. Scientific theories work together, so we cannot throw out one without undermining the others. That all these other scientific theories are not only well confirmed but useful is what makes scientists give up one principle rather than another when an anomaly arises.

At this point we might seek an even deeper explanation and ask why water expands when it freezes. In fact, to this day, nobody seems to have a fully adequate explanation of this common phenomenon. There are various theories but no agreement about how to explain the expansion of water. It has something to do with the way in which liquid water crystallizes to form ice, but nobody is quite sure why water crystallizes in that way. Does this show that certain phenomena are beyond scientific understanding? Probably not. But it does suggest that science may never be complete. More questions arise as science progresses, and there will always be questions that remain unanswered. As scientists discover and explain more and more phenomena and see connections among principles in different areas, every new step gives rise to more questions that need to be answered. That is one way in which science makes progress.

## SCIENTIFIC REVOLUTIONS

Another type of scientific development is more radical—knowledge is not simply extended, but, instead, one scientific framework is replaced (or largely replaced) by another. What changes is not just particular claims, but large-scale ways of doing science. In biology, the germ theory of disease and the theory of evolution through natural selection are examples of such revolutionary developments. Einstein's theory of relativity and the rise of quantum mechanics were also revolutionary developments in physics. Indeed, every branch of science has undergone at least one such revolutionary change during the past few centuries.

There are some important differences between scientific progress within a framework and the replacement of one framework by another.<sup>1</sup> In the first place, such changes in framework usually meet with strong resistance. A new conceptual framework will be unfamiliar and hard to understand and may even seem absurd or unintelligible. Even today, for example, the thought that the Earth is spinning on its axis and revolving around the sun seems completely counter to our commonsense view of the world. Also, arguments on behalf of a new framework will be very different from arguments that

occur *within* a framework. Disputes over conceptual frameworks cannot be settled by a straightforward appeal to facts. The long debate between Albert Einstein and Niels Bohr concerning quantum theory did not turn on matters of fact but on their interpretation. Einstein could not accept the indeterminacy involved in the quantum theory's interpretation of the world, and he worked until the end of his life to find some alternative to it. At present, few scientists share Einstein's reservations.

One of the most important and fascinating revolutions in science was set in motion by Darwin's theory of evolution through natural selection. The Darwinian revolution in biology challenged—and continues to challenge—many common assumptions about the nature of science and also the nature of humankind. As a result, Darwin's views on evolution and natural selection encountered vehement opposition right from the start.

These conflicts have attracted public attention when school boards have tried either to prevent the teaching of evolution in public schools or to require or allow alternative views, such as so-called "intelligent design," to be taught alongside standard evolutionary biology. In one recent case, the school board of Dover, Pennsylvania, passed a resolution requiring teachers to read a statement about "intelligent design" aloud in ninth-grade science classes whenever evolution was taught. Eleven parents of high-school students challenged this requirement in the case of *Kitzmiller v. Dover Area School District*, 400 F. Supp. 2d 707 (M.D. Pa. 2005). In his opinion, U.S. District Judge John E. Jones III ruled that intelligent design is not science, partly because it depends on "its creationist, and thus religious, antecedents," so the school district's requirement therefore violated the Establishment Clause of the First Amendment to the U.S. Constitution.

This debate between natural selection and intelligent design, thus, has not only scientific but also legal and religious dimensions. The following selections represent this debate at its best. Michael Behe (professor of biochemistry at Lehigh University in Pennsylvania) is one of the most prominent proponents of intelligent design, and Philip Kitcher (professor of philosophy at Columbia University) is one of his most profound critics. They focus on simple examples that are supposed to cause problems for the Darwinian approach to evolution. These case studies raise more general issues about the nature of religion, the nature of science, and the relation between religion and science—issues that remain alive today in pulpits, classrooms, and courts.

#### DISCUSSION QUESTION

To get simplistic objections out of the way first, explain what is wrong with Stephen Colbert's satirical refutation of Darwin's theory of evolution in the following passage:

Hey kids! Now you can disprove evolution in your own backyard. Here's what you'll need: one fishbowl, one pitcher of water, one hamster, alive, and one hardbound copy, Charles Darwin's *Origin of Species* . . . and now here's the experiment. . . .

- STEP 1:** Fill your fishbowl with the water. I don't want to give anything away, but soon it's going to be a bowl for another kind of animal.
- STEP 2:** Drop the hamster (you can call it "Skip") into the fishbowl.
- STEP 3:** Cover the fishbowl with Charles Darwin's *Origin of Species*.
- STEP 4:** Seems like a pretty desperate situation Skip has gotten himself into. This would be an ideal time for evolution to kick in.
- STEP 5:** Follow the scientific method—observe! Is the hamster "evolving" gills? Has he "evolved" a jackhammer to drill through the fishbowl, or "adapted to his environment" with a tiny hamster flamethrower to burn through *Origin of Species*? Don't think so.
- STEP 6:** Let the hamster go. Just because Darwin was a sick twist with a God complex doesn't mean we have to buy into his power trip. (You could also call the hamster "Teddy.")<sup>2</sup>

## MOLECULAR MACHINES: EXPERIMENTAL SUPPORT FOR THE DESIGN INFERENCE<sup>3</sup>

by Michael J. Behe

### DARWINISM'S PROSPERITY

Within a short time after Charles Darwin published *The Origin of Species* the explanatory power of the theory of evolution was recognized by the great majority of biologists. The hypothesis readily resolved the problems of homologous resemblance, rudimentary organs, species abundance, extinction, and biogeography. The rival theory of the time, which posited creation of species by a supernatural being, appeared to most reasonable minds to be much less plausible, since it would have a putative Creator attending to details that seemed to be beneath His dignity.

As time went on the theory of evolution obliterated the rival theory of creation, and virtually all working scientists studied the biological world from a Darwinian perspective. Most educated people now lived in a world where the wonder and diversity of the biological kingdom were produced by the simple, elegant principle of natural selection.

However, in science a successful theory is not necessarily a correct theory. In the course of history there have also been other theories which achieved the triumph that Darwinism achieved, which brought many experimental and observational facts into a coherent framework, and which appealed to people's intuitions about how the world should work. Those theories also promised to explain much of the universe with a few simple principles. But, by and large, those other theories are now dead.

A good example of this is the replacement of Newton's mechanical view of the universe by Einstein's relativistic universe. Although Newton's model

Source: Molecular Machines: Experimental Support for the Design Inference by Michael J. Behe from COSMIC PURSUIT, Vol. 1, No. 2, 1998, pp. 27–35. Reprinted by permission of the author.

accounted for the results of many experiments in his time, it failed to explain aspects of gravitation. Einstein solved that problem and others by completely rethinking the structure of the universe.

Similarly, Darwin's theory of evolution prospered by explaining much of the data of his time and the first half of the 20th century, but my article will show that Darwinism has been unable to account for phenomena uncovered by the efforts of modern biochemistry during the second half of this century. I will do this by emphasizing the fact that life at its most fundamental level is irreducibly complex and that such complexity is incompatible with undirected evolution.

#### A SERIES OF EYES

How do we see?

In the 19th century the anatomy of the eye was known in great detail and the sophisticated mechanisms it employs to deliver an accurate picture of the outside world astounded everyone who was familiar with them. Scientists of the 19th century correctly observed that if a person were so unfortunate as to be missing one of the eye's many integrated features, such as the lens, or iris, or ocular muscles, the inevitable result would be a severe loss of vision or outright blindness. Thus it was concluded that the eye could only function if it were nearly intact.

As Charles Darwin was considering possible objections to his theory of evolution by natural selection in *The Origin of Species* he discussed the problem of the eye in a section of the book appropriately entitled "Organs of extreme perfection and complication." He realized that if in one generation an organ of the complexity of the eye suddenly appeared, the event would be tantamount to a miracle. Somehow, for Darwinian evolution to be believable, the difficulty that the public had in envisioning the gradual formation of complex organs had to be removed.

Darwin succeeded brilliantly, not by actually describing a real pathway that evolution might have used in constructing the eye, but rather by pointing to a variety of animals that were known to have eyes of various constructions, ranging from a simple light sensitive spot to the complex vertebrate camera eye, and suggesting that the evolution of the human eye might have involved similar organs as intermediates.

But the question remains, how do we see? Although Darwin was able to persuade much of the world that a modern eye could be produced gradually from a much simpler structure, he did not even attempt to explain how the simple light sensitive spot that was his starting point actually worked. When discussing the eye Darwin dismissed the question of its ultimate mechanism by stating: "How a nerve comes to be sensitive to light hardly concerns us more than how life itself originated."

He had an excellent reason for declining to answer the question: 19th century science had not progressed to the point where the matter could even

be approached. The question of how the eye works—that is, what happens when a photon of light first impinges on the retina—simply could not be answered at that time. As a matter of fact, no question about the underlying mechanism of life could be answered at that time. How do animal muscles cause movement? How does photosynthesis work? How is energy extracted from food? How does the body fight infection? All such questions were unanswerable.

#### THE CALVIN AND HOBBS APPROACH

Now, it appears to be a characteristic of the human mind that when it lacks understanding of a process, then it seems easy to imagine simple steps leading from nonfunction to function. A happy example of this is seen in the popular comic strip *Calvin and Hobbes*. Little boy Calvin is always having adventures in the company of his tiger Hobbes by jumping in a box and traveling back in time, or grabbing a toy ray gun and “transmogrifying” himself into various animal shapes, or again using a box as a duplicator and making copies of himself to deal with worldly powers such as his mom and his teachers. A small child such as Calvin finds it easy to imagine that a box just might be able to fly like an airplane (or something), because Calvin doesn’t know how airplanes work.

A good example from the biological world of complex changes appearing to be simple is the belief in spontaneous generation. One of the chief proponents of the theory of spontaneous generation during the middle of the 19th century was Ernst Haeckel, a great admirer of Darwin and an eager popularizer of Darwin’s theory. From the limited view of cells that 19th century microscopes provided, Haeckel believed that a cell was a “simple little lump of albuminous combination of carbon,” not much different from a piece of microscopic Jell-O®. Thus it seemed to Haeckel that such simple life could easily be produced from inanimate material.

In 1859, the year of the publication of *The Origin of Species*, an exploratory vessel, the H.M.S. Cyclops, dredged up some curious-looking mud from the sea bottom. Eventually Haeckel came to observe the mud and thought that it closely resembled some cells he had seen under a microscope. Excitedly he brought this to the attention of no less a personage than Thomas Henry Huxley, Darwin’s great friend and defender, who observed the mud for himself. Huxley, too, became convinced that it was Urschleim (that is, protoplasm), the progenitor of life itself, and Huxley named the mud *Bathybius haeckelii* after the eminent proponent of abiogenesis.

The mud failed to grow. In later years, with the development of new biochemical techniques and improved microscopes, the complexity of the cell was revealed. The “simple lumps” were shown to contain thousands of different types of organic molecules, proteins, and nucleic acids, many discrete subcellular structures, specialized compartments for specialized processes, and an extremely complicated architecture. Looking back from

the perspective of our time, the episode of *Bathybius haeckelii* seems silly or downright embarrassing, but it shouldn't. Haeckel and Huxley were behaving naturally, like Calvin: since they were unaware of the complexity of cells, they found it easy to believe that cells could originate from simple mud.

Throughout history there have been many other examples, similar to that of Haeckel, Huxley, and the cell, where a key piece of a particular scientific puzzle was beyond the understanding of the age. In science there is even a whimsical term for a machine or structure or process that does something, but the actual mechanism by which it accomplishes its task is unknown: it is called a "black box." In Darwin's time all of biology was a black box: not only the cell, or the eye, or digestion, or immunity, but every biological structure and function because, ultimately, no one could explain how biological processes occurred.

Biology has progressed tremendously due to the model that Darwin put forth. But the black boxes Darwin accepted are now being opened, and our view of the world is again being shaken.

Take our modern understanding of proteins, for example.

#### PROTEINS

In order to understand the molecular basis of life it is necessary to understand how things called "proteins" work. Proteins are the machinery of living tissue that build the structures and carry out the chemical reactions necessary for life. For example, the first of many steps necessary for the conversion of sugar to biologically-useable forms of energy is carried out by a protein called hexokinase. Skin is made in large measure of a protein called collagen. When light impinges on your retina it interacts first with a protein called rhodopsin. A typical cell contains thousands and thousands of different types of proteins to perform the many tasks necessary for life, much like a carpenter's workshop might contain many different kinds of tools for various carpentry tasks.

What do these versatile tools look like? The basic structure of proteins is quite simple: they are formed by hooking together in a chain discrete sub-units called amino acids. Although the protein chain can consist of anywhere from about 50 to about 1,000 amino acid links, each position can only contain one of 20 different amino acids. In this they are much like words: words can come in various lengths but they are made up from a discrete set of 26 letters.

Now, a protein in a cell does not float around like a floppy chain; rather, it folds up into a very precise structure which can be quite different for different types of proteins. Two different amino acid sequences—two different proteins—can be folded to structures as specific and different from each other as a three-eighths inch wrench and a jigsaw. And like the household tools, if the shape of the proteins is significantly warped then they fail to do their jobs.

**THE EYESIGHT OF MAN**

In general, biological processes on the molecular level are performed by networks of proteins, each member of which carries out a particular task in a chain.

Let us return to the question, how do we see? Although to Darwin the primary event of vision was a black box, through the efforts of many biochemists an answer to the question of sight is at hand. The answer involves a long chain of steps that begin when light strikes the retina and a photon is absorbed by an organic molecule called 11-cis-retinal, causing it to rearrange itself within picoseconds. This causes a corresponding change to the protein, rhodopsin, which is tightly bound to it, so that it can react with another protein called transducin, which in turn causes a molecule called GDP to be exchanged with a molecule called GTP.

To make a long story short, this exchange begins a long series of further bindings between still more specialized molecular machinery, and scientists now understand a great deal about the system of gateways, pumps, ion channels, critical concentrations, and attenuated signals that result in a current to finally be transmitted down the optic nerve to the brain, interpreted as vision. Biochemists also understand the many chemical reactions involved in restoring all these changed or depleted parts to make a new cycle possible.

**TO EXPLAIN LIFE**

Although space doesn't permit me to give the details of the biochemistry of vision here, I have given the steps in my talks. Biochemists know what it means to "explain" vision. They know the level of explanation that biological science eventually must aim for. In order to say that some function is understood, every relevant step in the process must be elucidated. The relevant steps in biological processes occur ultimately at the molecular level, so a satisfactory explanation of a biological phenomenon such as sight, or digestion, or immunity, must include a molecular explanation.

It is no longer sufficient, now that the black box of vision has been opened, for an "evolutionary explanation" of that power to invoke only the anatomical structures of whole eyes, as Darwin did in the 19th century and as most popularizers of evolution continue to do today. Anatomy is, quite simply, irrelevant. So is the fossil record. It does not matter whether or not the fossil record is consistent with evolutionary theory, any more than it mattered in physics that Newton's theory was consistent with everyday experience. The fossil record has nothing to tell us about, say, whether or how the interactions of 11-cis-retinal with rhodopsin, transducin, and phosphodiesterase could have developed, step by step.

"How a nerve comes to be sensitive to light hardly concerns us more than how life itself originated," said Darwin in the 19th century. But both

phenomena have attracted the interest of modern biochemistry in the past few decades. The story of the slow paralysis of research on life's origin is quite interesting, but space precludes its retelling here. Suffice it to say that at present the field of origin-of-life studies has dissolved into a cacophony of conflicting models, each unconvincing, seriously incomplete, and incompatible with competing models. In private even most evolutionary biologists will admit that science has no explanation for the beginning of life.

The same problems which beset origin-of-life research also bedevil efforts to show how virtually any complex biochemical system came about. Biochemistry has revealed a molecular world which stoutly resists explanation by the same theory that has long been applied at the level of the whole organism. Neither of Darwin's black boxes—the origin of life or the origin of vision (or other complex biochemical systems)—has been accounted for by his theory.

#### IRREDUCIBLE COMPLEXITY

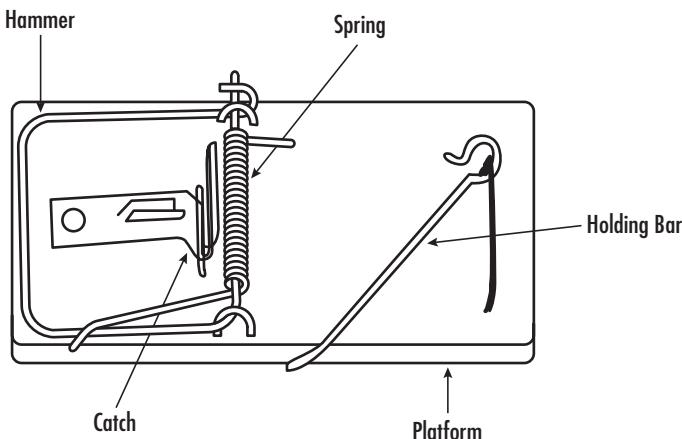
In *The Origin of Species* Darwin stated:

If it could be demonstrated that any complex organ existed which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down.

A system which meets Darwin's criterion is one which exhibits irreducible complexity. By irreducible complexity I mean a single system which is composed of several interacting parts that contribute to the basic function, and where the removal of any one of the parts causes the system to effectively cease functioning. An irreducibly complex system cannot be produced directly by slight, successive modification of a precursor system, since any precursor to an irreducibly complex system is by definition nonfunctional.

Since natural selection requires a function to select, an irreducibly complex biological system, if there is such a thing, would have to arise as an integrated unit for natural selection to have anything to act on. It is almost universally conceded that such a sudden event would be irreconcilable with the gradualism Darwin envisioned. At this point, however, "irreducibly complex" is just a term, whose power resides mostly in its definition. We must now ask if any real thing is in fact irreducibly complex, and, if so, then are any irreducibly complex things also biological systems?

Consider the humble mousetrap (Figure 20.1). The mousetraps that my family uses in our home to deal with unwelcome rodents consist of a number of parts. There are: (1) a flat wooden platform to act as a base; (2) a metal hammer, which does the actual job of crushing the little mouse; (3) a wire spring with extended ends to press against the platform and the hammer when the trap is charged; (4) a sensitive catch which releases when slight pressure is applied; and (5) a metal bar which holds the hammer back when the trap is charged and connects to the catch. There are also assorted staples and screws to hold the system together.



**Figure 20.1** A household mousetrap. The working parts of the trap are labeled. If any of the parts is missing, the trap does not function.

If any one of the components of the mousetrap (the base, hammer, spring, catch, or holding bar) is removed, then the trap does not function. In other words, the simple little mousetrap has no ability to trap a mouse until several separate parts are all assembled.

Because the mousetrap is necessarily composed of several parts, it is irreducibly complex. Thus, irreducibly complex systems exist.

#### MOLECULAR MACHINES

Now, are any biochemical systems irreducibly complex? Yes, it turns out that many are.

Earlier we discussed proteins. In many biological structures proteins are simply components of larger molecular machines. Like the picture tube, wires, metal bolts, and screws that comprise a television set, many proteins are part of structures that only function when virtually all of the components have been assembled.

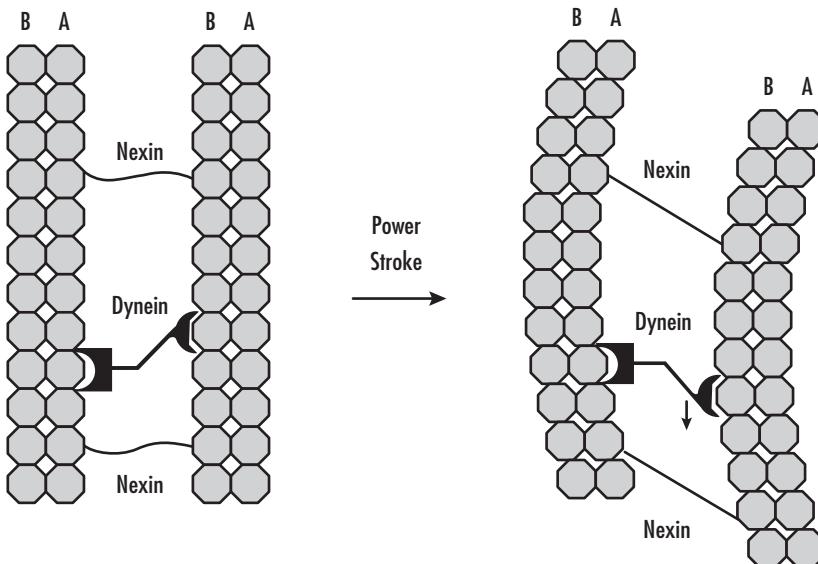
A good example of this is a cilium. Cilia are hairlike organelles on the surfaces of many animal and lower plant cells that serve to move fluid over the cell's surface or to "row" single cells through a fluid. In humans, for example, epithelial cells lining the respiratory tract each have about 200 cilia that beat in synchrony to sweep mucus towards the throat for elimination.

A cilium consists of a membrane-coated bundle of fibers called an axoneme. An axoneme contains a ring of 9 double microtubules surrounding two central single microtubules. Each outer doublet consists of a ring of 13 filaments (subfiber A) fused to an assembly of 10 filaments (subfiber B). The filaments of the microtubules are composed of two proteins called alpha and beta tubulin. The 11 microtubules forming an axoneme are held

together by three types of connectors: subfibers A are joined to the central microtubules by radial spokes; adjacent outer doublets are joined by linkers that consist of a highly elastic protein called nexin; and the central microtubules are joined by a connecting bridge. Finally, every subfiber A bears two arms, an inner arm and an outer arm, both containing the protein dynein.

But how does a cilium work? Experiments have indicated that ciliary motion results from the chemically-powered “walking” of the dynein arms on one microtubule up the neighboring subfiber B of a second microtubule so that the two microtubules slide past each other (Figure 20.2). However, the protein cross-links between microtubules in an intact cilium prevent neighboring microtubules from sliding past each other by more than a short distance. These cross-links, therefore, convert the dynein-induced sliding motion to a bending motion of the entire axoneme.

Now, let us sit back, review the workings of the cilium, and consider what it implies. Cilia are composed of at least a half dozen proteins: alpha-tubulin, beta-tubulin, dynein, nexin, spoke protein, and a central bridge protein. These combine to perform one task, ciliary motion, and all of these proteins must be present for the cilium to function. If the tubulins are absent, then there are no filaments to slide; if the dynein is missing, then the cilium remains rigid and motionless; if nexin or the other connecting proteins are missing, then the axoneme falls apart when the filaments slide.



**Figure 20.2** Schematic drawing of part of a cilium. The power stroke of the motor protein dynein, attached to one microtubule, against subfiber B of a neighboring microtubule causes the fibers to slide past each other. The flexible linker protein, nexin, converts the sliding motion to a bending motion.

What we see in the cilium, then, is not just profound complexity, but it is also irreducible complexity on the molecular scale. Recall that by “irreducible complexity” we mean an apparatus that requires several distinct components for the whole to work. My mousetrap must have a base, hammer, spring, catch, and holding bar, all working together, in order to function. Similarly, the cilium, as it is constituted, must have the sliding filaments, connecting proteins, and motor proteins for function to occur. In the absence of any one of those components, the apparatus is useless.

The components of cilia are single molecules. This means that there are no more black boxes to invoke; the complexity of the cilium is final, fundamental. And just as scientists, when they began to learn the complexities of the cell, realized how silly it was to think that life arose spontaneously in a single step or a few steps from ocean mud, so too we now realize that the complex cilium cannot be reached in a single step or a few steps.

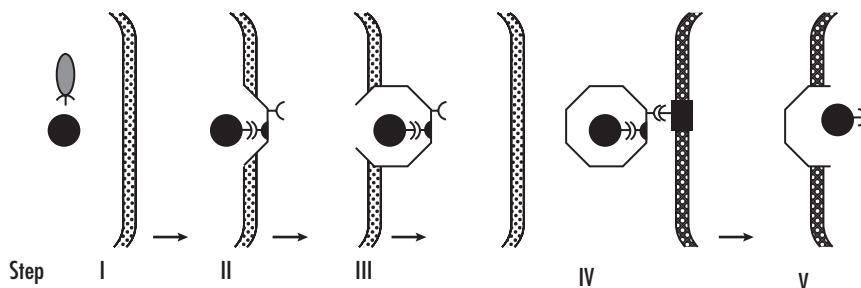
But since the complexity of the cilium is irreducible, then it cannot have functional precursors. Since the irreducibly complex cilium cannot have functional precursors, it cannot be produced by natural selection, which requires a continuum of function to work. Natural selection is powerless when there is no function to select. We can go further and say that, if the cilium cannot be produced by natural selection, then the cilium was designed.

#### A NON-MECHANICAL EXAMPLE

A non-mechanical example of irreducible complexity can be seen in the system that targets proteins for delivery to subcellular compartments. In order to find their way to the compartments where they are needed to perform specialized tasks, certain proteins contain a special amino acid sequence near the beginning called a “signal sequence.”

As the proteins are being synthesized by ribosomes, a complex molecular assemblage called the signal recognition particle or SRP, binds to the signal sequence. This causes synthesis of the protein to halt temporarily. During the pause in protein synthesis the SRP is bound by the transmembrane SRP receptor, which causes protein synthesis to resume and which allows passage of the protein into the interior of the endoplasmic reticulum (ER). As the protein passes into the ER the signal sequence is cut off.

For many proteins the ER is just a way station on their travels to their final destinations (Figure 20.3). Proteins which will end up in a lysosome are enzymatically “tagged” with a carbohydrate residue called mannose-6-phosphate while still in the ER. An area of the ER membrane then begins to concentrate several proteins; one protein, clathrin, forms a sort of geodesic dome called a coated vesicle which buds off from the ER. In the dome there is also a receptor protein which binds to both the clathrin and to the mannose-6-phosphate group of the protein which is being transported.



**Figure 20.3** Transport of a protein from the ER to the lysosome. Step I: A specific enzyme (gray oval) places a marker on the protein (black sphere). This takes place within the ER, which is delimited by a barrier membrane (cross-hatched bar with ends curving to the left). Step II: The marker is specifically recognized by a receptor protein and the clathrin vesicle (hexagonal shape) begins to form. Step III: The clathrin vesicle is completed and buds off from the ER membrane. Step IV: The clathrin vesicle crosses the cytoplasm and attaches through another specific marker to a receptor protein (dark gray box) on the lysosomal membrane and releases its cargo.

The coated vesicle then leaves the ER, travels through the cytoplasm, and binds to the lysosome through another specific receptor protein. Finally, in a maneuver involving several more proteins, the vesicle fuses with the lysosome and the protein arrives at its destination.

During its travels our protein interacted with dozens of macromolecules to achieve one purpose: its arrival in the lysosome. Virtually all components of the transport system are necessary for the system to operate, and therefore the system is irreducible. And since all of the components of the system are comprised of single or several molecules, there are no black boxes to invoke. The consequences of even a single gap in the transport chain can be seen in the hereditary defect known as I-cell disease. It results from a deficiency of the enzyme that places the mannose-6-phosphate on proteins to be targeted to the lysosomes. I-cell disease is characterized by progressive retardation, skeletal deformities, and early death.

#### THE STUDY OF “MOLECULAR EVOLUTION”

Other examples of irreducible complexity abound, including aspects of protein transport, blood clotting, closed circular DNA, electron transport, the bacterial flagellum, telomeres, photosynthesis, transcription regulation, and much more. Examples of irreducible complexity can be found on virtually every page of a biochemistry textbook. But if these things cannot be explained by Darwinian evolution, how has the scientific community regarded these phenomena of the past forty years?

A good place to look for an answer to that question is in the *Journal of Molecular Evolution*. *JME* is a journal that was begun specifically to deal with the topic of how evolution occurs on the molecular level. It has high scientific standards, and is edited by prominent figures in the field. In a recent issue of *JME* there were published eleven articles; of these, all eleven were concerned simply with the analysis of protein or DNA sequences. None of the papers discussed detailed models for intermediates in the development of complex biomolecular structures. In the past ten years *JME* has published 886 papers. Of these, 95 discussed the chemical synthesis of molecules thought to be necessary for the origin of life, 44 proposed mathematical models to improve sequence analysis, 20 concerned the evolutionary implications of current structures and 719 were analyses of protein or polynucleotide sequences. However, there weren't any papers discussing detailed models for intermediates in the development of complex biomolecular structures. This is not a peculiarity of *JME*. No papers are to be found that discuss detailed models for intermediates in the development of complex biomolecular structures in the *Proceedings of the National Academy of Science*, *Nature*, *Science*, the *Journal of Molecular Biology* or, to my knowledge, any journal whatsoever.

Sequence comparisons overwhelmingly dominate the literature of molecular evolution. But sequence comparisons simply can't account for the development of complex biochemical systems any more than Darwin's comparison of simple and complex eyes told him how vision worked. Thus in this area science is mute.

#### DETECTION OF DESIGN

What's going on? Imagine a room in which a body lies crushed, flat as a pancake. A dozen detectives crawl around, examining the floor with magnifying glasses for any clue to the identity of the perpetrator. In the middle of the room next to the body stands a large, gray elephant. The detectives carefully avoid bumping into the pachyderm's legs as they crawl, and never even glance at it. Over time the detectives get frustrated with their lack of progress but resolutely press on, looking even more closely at the floor. You see, textbooks say detectives must "get their man," so they never consider elephants.

There is an elephant in the roomful of scientists who are trying to explain the development of life. The elephant is labeled "intelligent design." To a person who does not feel obliged to restrict his search to unintelligent causes, the straightforward conclusion is that many biochemical systems were designed. They were designed not by the laws of nature, not by chance and necessity. Rather, they were planned. The designer knew what the systems would look like when they were completed; the designer took steps to bring the systems about. Life on earth

at its most fundamental level, in its most critical components, is the product of intelligent activity.

The conclusion of intelligent design flows naturally from the data itself—not from sacred books or sectarian beliefs. Inferring that biochemical systems were designed by an intelligent agent is a humdrum process that requires no new principles of logic or science. It comes simply from the hard work that biochemistry has done over the past forty years, combined with consideration of the way in which we reach conclusions of design every day.

What is “design”? Design is simply the purposeful arrangement of parts. The scientific question is how we detect design. This can be done in various ways, but design can most easily be inferred for mechanical objects.

Systems made entirely from natural components can also evince design. For example, suppose you are walking with a friend in the woods. All of a sudden your friend is pulled high in the air and left dangling by his foot from a vine attached to a tree branch.

After cutting him down you reconstruct the trap. You see that the vine was wrapped around the tree branch, and the end pulled tightly down to the ground. It was securely anchored to the ground by a forked branch. The branch was attached to another vine—hidden by leaves—so that, when the trigger-vine was disturbed, it would pull down the forked stick, releasing the spring-vine. The end of the vine formed a loop with a slipknot to grab an appendage and snap it up into the air. Even though the trap was made completely of natural materials you would quickly conclude that it was the product of intelligent design.

Intelligent design is a good explanation for a number of biochemical systems, but I should insert a word of caution. Intelligent design theory has to be seen in context: it does not try to explain everything. We live in a complex world where lots of different things can happen. When deciding how various rocks came to be shaped the way they are a geologist might consider a whole range of factors: rain, wind, the movement of glaciers, the activity of moss and lichens, volcanic action, nuclear explosions, asteroid impact, or the hand of a sculptor. The shape of one rock might have been determined primarily by one mechanism, the shape of another rock by another mechanism.

Similarly, evolutionary biologists have recognized that a number of factors might have affected the development of life: common descent, natural selection, migration, population size, founder effects (effects that may be due to the limited number of organisms that begin a new species), genetic drift (spread of “neutral,” nonselective mutations), gene flow (the incorporation of genes into a population from a separate population), linkage (occurrence of two genes on the same chromosome), and much more. The fact that some biochemical systems were designed by an intelligent agent does not mean that any of the other factors are not operative, common, or important.

**CONCLUSION**

It is often said that science must avoid any conclusions which smack of the supernatural. But this seems to me to be both bad logic and bad science. Science is not a game in which arbitrary rules are used to decide what explanations are to be permitted. Rather, it is an effort to make true statements about physical reality. It was only about sixty years ago that the expansion of the universe was first observed. This fact immediately suggested a singular event—that at some time in the distant past the universe began expanding from an extremely small size.

To many people this inference was loaded with overtones of a supernatural event—the creation, the beginning of the universe. The prominent physicist A. S. Eddington probably spoke for many physicists in voicing his disgust with such a notion:

Philosophically, the notion of an abrupt beginning to the present order of Nature is repugnant to me, as I think it must be to most; and even those who would welcome a proof of the intervention of a Creator will probably consider that a single winding-up at some remote epoch is not really the kind of relation between God and his world that brings satisfaction to the mind.

Nonetheless, the big bang hypothesis was embraced by physics and over the years has proven to be a very fruitful paradigm. The point here is that physics followed the data where it seemed to lead, even though some thought the model gave aid and comfort to religion. In the present day, as biochemistry multiplies examples of fantastically complex molecular systems, systems which discourage even an attempt to explain how they may have arisen, we should take a lesson from physics. The conclusion of design flows naturally from the data; we should not shrink from it; we should embrace it and build on it.

In concluding, it is important to realize that we are not inferring design from what we do not know, but from what we do know. We are not inferring design to account for a black box, but to account for an open box. A man from a primitive culture who sees an automobile might guess that it was powered by the wind or by an antelope hidden under the car, but when he opens up the hood and sees the engine he immediately realizes that it was designed. In the same way biochemistry has opened up the cell to examine what makes it run and we see that it, too, was designed.

It was a shock to the people of the 19th century when they discovered, from observations science had made, that many features of the biological world could be ascribed to the elegant principle of natural selection. It is a shock to us in the twentieth century to discover, from observations science has made, that the fundamental mechanisms of life cannot be ascribed to natural selection, and therefore were designed. But we must deal with our shock as best we can and go on. The theory of undirected evolution is already dead, but the work of science continues.

**DISCUSSION QUESTIONS**

1. Explain what Behe means by “irreducible complexity.” Give examples of what is and of what is not irreducibly complex, according to Behe, then explain why he takes the former to be and the latter not to be irreducibly complex. Why is irreducible complexity supposed to be a problem for Darwin’s theory of evolution?
2. Behe cites the human eye as evidence against Darwin’s theory of evolution (compare also Paley on p. 213). Explain Behe’s claims and arguments using this example. How could Darwin best respond? Do you think Darwin’s theory of evolution can explain the eye? Why or why not?
3. Behe next cites cilia to argue against Darwin’s theory of evolution. Answer the same questions as in Discussion Question 2 about this example.
4. In addition to arguing against Darwin, Behe argues positively that “the conclusion of intelligent design flows naturally from the data itself—not from sacred books or sectarian beliefs.” Explain Behe’s conclusion here and his argument for that conclusion. Do you agree with Behe? Why or why not?

**LIVING WITH DARWIN<sup>4</sup>**

by Philip Kitcher

... Darwin’s own consideration of the concrete case argument focused on some complex organs and structures that he rightly believed to be hard to understand in terms of natural selection. Two examples are prominent in the *Origin*, the eye and the electric organs found in some fish. The latter example disconcerted Darwin because the fish with electric organs are of very different types, and have their organs in different parts of their bodies. Much to his relief, research on electric fish carried out by a contemporary who was not sympathetic to evolutionary ideas—“McDonnell of Dublin (a first-rate man)”\*—revealed that, for each type of fish with an electric organ, there is a related fish with a similar organ (not functionally electric) in the same position. What had initially appeared to be the challenge of understanding how different electric organs had been built from scratch, became the much simpler question of how a similar change had occurred in each instance.

Darwin himself offered a tentative proposal about the evolution of the eye. He supposed that sensitivity to light might come in degrees, and that it might be possible to find, among existing organisms, some with a crude ability to respond to light, others with a more refined capacity and so on in something like a series. Perhaps, he speculated, research on these creatures might expose reasons why the different levels of sensitivity provided

\* Francis Darwin, ed., *Life and Letters of Charles Darwin*, vol. 2 (London: John Murray, 1888), 352.

an advantage over rival organisms who had less, thus providing a way of answering (or sidestepping) the creationist quip, "What use is half an eye?"

It has taken more than a century of research on a wide variety of organisms to demonstrate that Darwin's hunch was basically right. Appearances to the contrary, organs and structures sensitive to light can be assembled piecemeal, with the intermediates enjoying some advantage over the competition. Biologists have studied organisms that respond to the light that impinges on their surfaces, organisms with indentations of the superficial layer that are able to acquire information about the direction of the light, organisms with deeper indentations whose light detection is more fine grained, organisms that have a structure resembling a pinhole camera, organisms that interpose a translucent medium between the surface and the aperture through which the light comes—and so on. By studying this sequence of organisms, they have been able to explore the transitions through which relatively crude abilities to detect light were successively refined.\* One feature of the story deserves emphasis. Darwin didn't start with a comparison between the fully formed eye—in a human being or an octopus, say—and then think of the component parts as being introduced, one at a time. He resisted the challenge to explain first the advantage of an eighth of an eye, then the advantage of a quarter of an eye, and so on, and focused instead on a function, light sensitivity, that might have been refined from an initial state of absence. To put it more bluntly, he didn't allow his envisaged challengers to define the sequence of "intermediates" for him.

Savvy champions of the concrete case argument know this story. They appreciate Darwin's ingenuity in responding to the challenge, and, although they think the response ultimately fails, their reasons for this judgment depend on a more general problem for evolution under natural selection. That more general problem derives from the fine structure of the components of complex organs (like eyes), the molecular mechanisms that have to be in place for eyes to work. For all Darwin's cleverness, he failed to appreciate the full depth, and the full generality of the difficulty confronting him.

The principal exponent of the complex case argument is Michael Behe, a professor of biochemistry, who argues at length in *Darwin's Black Box* that the real troubles of natural selection become visible when you appreciate the molecular components of complex biological systems. Almost everywhere you look in nature, there are complicated structures and processes, with many molecular constituents, and all the constituents need to be present and to fit together precisely for things to work as they should. Biochemical pathways require numerous enzymes to interact with one another, in appropriate relative concentrations, so that some important process can occur. If you imagine a mutation in one of the genes that directs the formation of some

\* For a superbly accessible presentation of this research, see Richard Dawkins, *Climbing Mount Improbable* (New York: Norton, 1996). Another lucid, and concise, account by one of the scientists involved is Dan-Eric Nilsson, "Vision Optics and Evolution," *Bioscience* 39 (1989): 298–307.

essential protein, or if you suppose that the genetic material becomes shuffled in a way that allows for differences in the rates at which proteins are formed, it looks as though disaster will ensue. Crucial pieces will be missing, or won't be present in the right proportions, so that everything will break down. How then could organisms with the pertinent structures or processes have evolved from organisms that lacked them?

Behe offers numerous instances of molecular machines that, he claims, could not have been built up in stages by natural selection. Among his most influential examples is a discussion of devices that some bacteria use for motion, flagella. He contrasts the bacterial flagellum with a different motor, used by other cells, the cilium. "In 1973 it was discovered that some bacteria swim by rotating their flagella. So the bacterial flagellum acts as a rotary propeller—in contrast to the cilium, which acts more like an oar."<sup>\*</sup> Both flagella and cilia are intricate structures, and Behe describes the many molecular parts and systems that have to be present if they are to do their jobs. He concludes that the complexity of the organization dooms any attempt to explain its emergence as the result of natural selection. "As biochemists have begun to examine apparently simple structures like cilia and flagella, they have discovered staggering complexity, with dozens or even hundreds of precisely tailored parts. It is very likely that many of the parts we have not considered here are required for any cilium to function in a cell. As the number of required parts increases, the difficulty of gradually putting the system together skyrockets, and the likelihood of indirect scenarios plummets. Darwin looks more and more forlorn."<sup>\*\*</sup> Indeed, the most famous portraits of Darwin hardly make him look exactly cheerful, but it's worth asking why examples like these should render him more forlorn.

Perhaps it seems obvious. Natural selection depends upon mutations that are not produced in response to the organism's needs. The bacteria are at the mercy of chance, which will fling in this variant protein or that, with negligible probability that the latest novelty will fit with what went before or will contribute to the design project of building a flagellum. In essentials, however, this is precisely parallel to an old creationist strategy, just the one that Darwin sidestepped in the case of the eye. Behe has specified how the intermediates are to be formed, and it isn't surprising that his preferred scenario has the air of impossibility.

What exactly is known about the bacterial flagellum? During the past few decades, careful molecular studies have identified the genes that direct the assembly of the motor, and have explored the ways in which it is put

\* Behe, *Darwin's Black Box*, 70. Behe uses the example of the bacterial flagellum as a parade case in many of his writings and presentations. See, for example, "Design at the Foundation of Life," in *Science and Evidence for Design in the Universe* (San Francisco, CA: Ignatius Press, 2000), 120 ff.

\*\* Behe, *Darwin's Black Box*, 73.

together in the development of an individual bacterium. Some mutations in these genes allow for bacteria to move, albeit less efficiently. What is currently missing, however, is that systematic study of the differences among bacteria with flagella and bacteria without that would parallel the knowledge attained in the case of vision. A sufficiently intensive study of the genomes of bacteria that lack flagella would enable biologists to explore the potential role of some of the crucial genes, and of the proteins they give rise to, when others are absent, and thus enable them to make more progress with Behe's apparently formidable challenge.

Most sciences face unsolved problems—indeed the exciting unsolved problems are the motivators for talented people to enter a field. Chemists still struggle to understand how newly made proteins fold into their three-dimensional shapes as they are synthesized. To take an instance closer to hand, Behe's own discussion acknowledges that there's still a lot to learn about the molecular structure and functions of cilia. Unsolved questions are not typically written off as unsolvable—nobody proposes that there's some special force, unknown to current chemistry (an "intelligent force" perhaps?) that guides the proteins to their proper forms, or some hand that assembles the cilium in the development of an individual bacterium. Why, then, should we believe that the problem of the bacterial flagellum is unsolvable? Just because, in the absence of systematic molecular studies of bacteria with and without flagella, we can't currently give a satisfactory scenario for the evolution of the bacterial flagellum under natural selection, why should we conclude that further research couldn't disclose how that evolution occurred?

We are beguiled by the simple story line Behe rehearses. He invites us to consider the situation by supposing that the flagellum requires the introduction of some number—20, say—of proteins that the ancestral bacterium doesn't originally have. So Darwinians have to produce a sequence of 21 organisms, the first having none of the proteins, and each subsequent organism having one more than its predecessor. Darwin is forlorn because however he tries to imagine the possible pathway along which genetic changes successively appeared, he appreciates the plight of numbers 2–20, each of which is clogged with proteins that can't serve any function, proteins that interfere with important cellular processes. These organisms will be targets of selection, and will wither in the struggle for existence. Only number 1, and number 21, in which all the protein constituents come together to form the flagellum, have what it takes. Because of the dreadful plight of the intermediates, natural selection couldn't have brought the bacterium from there to here.

The story is fantasy, and Darwinians should disavow any commitment to it. First, there is no good reason for supposing that the ancestral bacterium lacked all, or even any, of the proteins needed to build the flagellum. It's a common theme of evolutionary biology that constituents of a cell, a tissue, or an organism are put to new uses because of a modification of the genome.

Perhaps the immediate precursor of the bacterium with the flagellum is an organism in which all the protein constituents are already present, but are employed in different ways. Then, at the very last step there's a change in the genome that removes whatever chemical barrier previously prevented the building of the flagellum. In this organism (the precursor), the function of one of the proteins is to increase the efficiency of a particular energy-transfer process. The precursor of the precursor lacked that protein, so that the genetic change that led to the precursor improved a process that was previously adequate. So it goes, back down a sequence of ancestors, all quite capable of functioning in their environments but all at a selective disadvantage to the bacteria that succeeded them.

Isn't this all fantasy too? Of course\*—but it is no more the product of speculative imagination than Behe's seemingly plausible assumption that the components of the flagellum would have had to be added one by one, and would have sat around idly (at best) until the culminating moment when all were present. Moreover, we were supposed to be offered a proof of impossibility, and that won't be complete until Behe and his allies have shown that all the conceivable scenarios through which bacteria might acquire flagella are flawed. Really demonstrating impossibility—or even improbability—here and in kindred instances, is extremely difficult, precisely because it would require a much more systematic survey of the molecular differences among bacteria.

The serious way forward is to amend our ignorance, by sequencing the genomes of different bacteria, with and without flagella. Using our current knowledge of the genetic basis of the flagellum, researchers would be able to specify more clearly what the intermediate forms—those with some, but not all, of the crucial genes—might have been like, and what functions the relevant proteins might have served. Until we know these things, efforts to describe intermediates will be so much whistling in the dark. Behe's examples rely on guesses that simply anticipate what this hard work would reveal.

So we have the illusion of an impossibility proof. Allegedly there could be no sequence of intermediates concluding with the fortunate, flagellum-bearing bacterium, in which each member of the sequence enjoyed a selective advantage over its predecessor. Behe's story (quite charmingly told in *Darwin's Black Box*) offers his own preferred version of what the sequence would have to be like. Since Darwinians have no commitment to simpleminded stories of sequential addition of components, there is no reason to accept Behe's description. Because the same rhetorical strategy pervades his entire book, showing up in all the instances of the concrete case argument he provides, all the

\* Perhaps not complete fantasy. The account I offer here is concordant with a recent review of the molecular details of the bacterial flagellum. See Howard C. Berg, "The Rotary Motor of Bacterial Flagella," *Annual Review of Biochemistry* 72 (2003): 19–54. (For this reference, I'm indebted to Mel Simon.)

parade of examples really shows is that there are some interesting problems for molecularly minded evolutionists to work on, problems they might hope to solve in the light of increased understanding from comparative studies of the genetics and development of a wide variety of organisms. . . .

So far I have focused on the negative doctrine of intelligent design, the identification of unsolved evolutionary problems. We now have to consider the positive thesis, the claim that the phenomena to which Darwin's detractors point are produced by a process that deserves the label "intelligent." Two issues need to be considered. First, on what grounds should we apply the label? Second, what help can intelligent design provide in understanding the phenomena in question? . . .

Making any judgment about whether a mechanism is intelligent or not appears rather difficult until we have been told considerably more about the way in which that mechanism operates. Officially, of course, we aren't supposed to personify this mechanism, and it's hard to understand just what the attribution of Intelligence even means if we resist the personification. If something counts as intelligent, wouldn't it have psychological states and engage in psychological processes—and wouldn't anything like that be very like a person? Intelligent design-ers do not address such questions. . . .

It's simply a fallacy to suppose that because a particular structure or mechanism appears complex, then the causal agent that brought it about must be appropriately characterized as having "foreseen" or "planned" or "designed" the outcome. Even if intelligent design-ers were right in supposing that the phenomena they indicate couldn't have evolved by natural selection, only a more explicit identification of the causal mechanism that was at work could justify the conclusion that that mechanism is intelligent.

So, turning to the second question posed above, what help can intelligent design provide when we try to understand the difficulties it takes to beset Darwinism? How does it deal with the bacterial flagellum, for example?

If we take Behe at his word when he declares that he finds "the idea of common descent" to be "fairly convincing," and that he has "no particular reason to doubt it,"\* then we should suppose that bacteria with flagella emerged from ancestors who lacked flagella. In line with the simple additive story he uses to make a history of natural selection appear implausible, he must suppose that the ancestors were missing a number of crucial proteins that the lucky descendants acquired, proteins that, once present, fit themselves together in the flagellum. If the intelligent design perspective is to help settle the unsolved problems of evolution, it would be good to have an alternative account that tells us how Intelligence facilitated the transition.

Unfortunately, the rest is silence. Neither in Behe's writings, nor in those of any other intelligent design-er, is there the slightest indication of how Intelligence performs the magic that poor, limited, natural selection cannot.

\* Behe, *Darwin's Black Box*, 7.

On the face of it, there are just two basic possibilities. The first option is that Intelligence arranges the environment so that the intermediates—the apparently hapless organisms, cluttered with useless proteins—are protected against elimination under natural selection. (If we were unofficially inclined, we might say that the good Lord tempers the wind to the shorn bacterium.) The second option is that Intelligence provides for coordinated mutations to arise. If 20 genetic changes are needed, it brings about all of them at once. Or we can mix elements of both options and suppose that Intelligence introduces mutations in clumps—first ten, say, and then another ten, or first seven, then another seven, then six—and arranges protective environments for the intermediates. Of course, any story along these lines raises serious doubts. Just how does the coordination of the genetic changes or the modification of the environment work?

In presenting the possibilities in this way, I may seem to be forcing words into the mouths of the intelligent design-ers. Their core position, after all, is that at crucial moments in the history of life, descendants of some ancestors who lacked some trait (or organ or structure) came to possess the pertinent trait (organ, structure) by some causal process that is, unlike natural selection, intelligent. Why, then, do they have to talk about genes, mutations, and the need for protection against natural selection? The answer is that the traits in question are heritable—they are not introduced in each generation by some continued activity on the part of Intelligence, but emerge through the interactions of genes and environments. As in the case of the bacterial flagellum, there are underlying genes, and hence there have to be genetic changes in the passages from the ancestors to the descendants. If these changes occur over several generations, then, on the intelligent design-ers' own principles, there has to be protection against the tendency of natural selection to weed out the hapless intermediates. If they happen in one step, then, again by the favored principles, there must be coordinated mutations. Hence, even if the position would prefer to talk more vaguely of “novelties,” it is committed to one of the options I have presented.

What intelligent design urgently needs if it's going to make any progress in understanding these transitions, in tackling the problems it claims to raise, is a set of coherent principles that identify the ways in which Intelligence is directed and what its powers and limitations are. If we lapse from the official story for a moment, we have to have some idea about what Intelligence “wants to achieve” and what kinds of things “it can do to work toward what it wants.” What basis do we have to think that Intelligence aims to remedy the plight of the flagellumless bacteria, who can't evolve into bacteria-with-a-flagellum under natural selection? What basis is there to believe that Intelligence—or anything else, for that matter—can coordinate genetic changes or modify environments?

In fact, we need two distinct kinds of principles. First, there have to be principles that specify when Intelligence swings into action. Perhaps they

will tell us that Intelligence operates when there are potentially advantageous complex traits that can't evolve by natural selection. Second, there must be principles that explain what Intelligence does when it acts. Perhaps these will identify the sorts of genetic changes Intelligence can arrange, or the ways in which it can inhibit the normal operation of selection.

It is already clear that these principles will be hard to state precisely. For, if Intelligence has been waiting in the wings throughout the history of life, seizing opportunities as they arise, we know that there are all sorts of things it hasn't done. Apparently Intelligence isn't directed toward eliminating the junk from genomes or removing vestigial structures like the whale's pelvis or generating radically new arrangements for mammalian forelimbs. It's possible, of course, that although directed toward these ends, Intelligence is simply unable to bring them about. So any satisfactory principles must differentiate between the bacterial flagellum, blood-clotting cascade, and similar places where Intelligence shows its prowess, and the accumulated junk, vestigial structures, and genetic blunders, where it remains in abeyance. . . .

Why do intelligent designers ignore the basic problem of explaining the power and direction of the mechanism they invoke, a problem that strikes at the heart of their theory? Apparently, their preferred perspective faces a multitude of currently unsolved puzzles about the scope and direction of Intelligence. Yet, unlike their counterparts in other scientific ventures, they are reluctant to suggest their strategies for seeking solutions. Their reticence provokes the charge that what they are doing is not science, but perhaps breaking their silence would be theologically unwise. Saying too much might disrupt the harmony between the sanitized version of intelligent design elaborated in the classroom during the week and the richer account delivered from the pulpit on Sunday. Moreover, saying anything that would genuinely respond to the puzzles might be saying too much.

Yet, I suspect many people would simply reject the terms in which I have posed the problem. Friends of intelligent design would prefer not to talk about evolutionary transitions at all. So, they might say, the complex structures are built from scratch. Intelligence is a creative force that replaces older types of organisms with new, individually designed species. Conceived in this way, intelligent design disavows Behe's acceptance of descent with modification, drawing bars across the tree of life to mark the places of radical discontinuity, of events of special creation. . . .

The real situation is that intelligent design-ers oscillate. . . . Yet however they wriggle, they find no satisfactory positive doctrine, no set of principles about Intelligence that can adequately account for the phenomena. This is why readers hunt through their literature seeking fragments of positive theory in vain. They won't find it. Because to advance any such theory would expose the corpse of dead science.

**DISCUSSION QUESTIONS**

1. How does Kitcher explain the evolution of the human eye? Is this explanation adequate? Why or why not?
2. How does Kitcher explain the evolution of the bacterial flagellum? Is this explanation adequate? Why or why not?
3. Are there examples that Darwinians cannot explain? What does this show?
4. Can intelligent design theory itself explain the examples that are supposed to cause trouble for Darwinian theories of evolution?
5. Is “intelligent design” a scientific theory? Why or why not?
6. In your opinion, should “intelligent design” be required (or allowed) to be taught in public high schools alongside evolution and natural selection as part of the science curriculum? Why or why not? (It might help to read Judge Jones’s opinion in *Kitzmiller v. Dover Area School District*.)

**NOTES**

<sup>1</sup> Thomas Kuhn gives prominence to this difference in his important work, *The Structure of Scientific Revolutions*, 2nd ed. (Chicago, IL: University of Chicago Press, 1970).

<sup>2</sup> Stephen Colbert, *I Am America (And So Can You!)* (New York: Grand Central Publishing, 2007), 207.

<sup>3</sup> From *Cosmic Pursuit* 1, no. 2 (1998): 27–35.

<sup>4</sup> *Living with Darwin* (New York: Oxford University Press, 2007), 85–95.

## RELIGIOUS REASONING

**R**eligion is central to the lives of many people across the world. They attend religious ceremonies and pray or meditate, sometimes several times a day. Even when they are not engaged in explicitly religious practices, religion also affects believers' views in morality, politics, and even science. (See Chapter 20.) Critics of religion disdain these influences and claim that religions depend on false or unjustified beliefs, especially the belief that God exists. Both sides of this debate present ingenious arguments, which will be explored in this chapter.

What is religion? That's not an easy question to answer. Religions are varied and complex. They include rituals, communities, institutions, texts, and beliefs. You cannot understand religion as a whole by looking merely at religious beliefs. Moreover, religious beliefs cover many topics, from the afterlife and the meaning of life to creation and miracles. You also cannot understand religion by considering only beliefs *about God*. Nonetheless, since belief in God is so central to so many religions, it is crucial to ask whether God exists.

People who believe that God exists are called *theists*. People who deny the existence of God (or any other god) are called *atheists*. Many people accept neither belief and claim that we cannot know whether God exists. They are called *agnostics*.

It is also important to realize, however, that people who believe in a traditional God within Christianity, say, usually deny the existence of other kinds of gods. Their position is, thus, theism with respect to their own religion and God but atheism regarding other religions and gods. Hence, when asking whether someone is a theist or an atheist, it is crucial to specify the kind of God at issue.

Who or what is God? Different religions include different beliefs about God or gods. Nonetheless, the dominant traditions in the Western world—Judaism, Christianity, and Islam—all derive from Abraham and share certain central beliefs about God. In traditional views, God is a person with the three “omnis”: omnipotent (or all-powerful), omniscient (or all-knowing), and omnibenevolent (or all-good). Some theologians add that God is omnipresent (or equally present at all spaces and times), but many see God as eternal or existing outside of all time and space. Nonetheless, God is supposed to be active in time, such as when He causes miracles or answers prayers. Not all

Jews, Christians, and Moslems share all of these claims about God, but they are accepted by traditional theologians.

The readings in this chapter are about the traditional Christian God. William Lane Craig argues for the existence of a traditional Christian God. Edwin Curley then argues against the existence of such a God. These readings are slightly revised transcripts of a live debate that occurred at the University of Michigan in 1999, so they bear some marks of that setting.

## FIVE REASONS TO BELIEVE IN GOD

by William Lane Craig

Good Evening! I want to begin by thanking [Michigan Christian Grads] for inviting me to participate in tonight's debate. And I want to say what a privilege it is to be debating so eminent a scholar as Professor Curley. When I was a doctoral student writing my dissertation on the cosmological argument for God's existence, Dr. Curley's work on the famous philosopher Benedict de Spinoza was a valuable resource to me in trying to analyze Spinoza's own argument for God. So it's a genuine honor to be sharing the podium with Dr. Curley tonight.

Now in tonight's debate it seems that there are two basic questions that we need to ask ourselves:

I. Are there any good reasons to think that God does not exist?

And

II. Are there good reasons to think that God does exist?

Now with respect to the first question, I'll leave it up to Dr. Curley to present the reasons why he thinks that God does not exist. Atheist philosophers have tried for centuries to disprove the existence of God. But no one has ever been able to come up with a convincing argument. So rather than attack straw men at this point, I'll just wait to hear Professor Curley's answer to the following question: What good reasons are there to think that God does not exist?

So let's move on, then, to that second question: Are there good reasons to think that God does exist? Tonight I'm going to present five reasons why I think that God exists. Whole books have been written on each one of these, so all I can present here is a brief sketch of each argument and then go into more detail as Dr. Curley responds to them.\* These reasons are independent of one another, so that if even one of them is sound, it furnishes good grounds for believing that God exists. Taken together, they constitute a powerful cumulative case that God exists.

\* For a popular presentation of these arguments and responses to typical objections, see my booklet *God, Are You There?* (Atlanta: RZIM, 1999).

Source: Five Reasons to Believe in God by William Lane Craig from The Craig-Curley Debate: The Existence of the Christian God from Reasonable Faith, 2007. Reprinted by permission of William Lane Craig.

1: God makes sense of the origin of the universe. Have you ever asked yourself where the universe came from? Why everything exists instead of just nothing? Typically atheists have said that the universe is eternal, and that's all. But surely this doesn't make sense. Just think about it for a minute. If the universe never began to exist, then that means that the number of events in the past history of the universe is infinite. But mathematicians recognize that the idea of an actually infinite number of things leads to self-contradictions. For example, what is infinity minus infinity? Well, mathematically, you get self-contradictory answers. This shows that infinity is just an idea in your mind, not something that exists in reality. David Hilbert, perhaps the greatest mathematician of this century states, "The infinite is nowhere to be found in reality. It neither exists in nature nor provides a legitimate basis for rational thought. The role that remains for the infinite to play is solely that of an idea."<sup>\*</sup>

But that entails that since past events are not just ideas, but are real, the number of past events must be finite. Therefore, the series of past events can't just go back forever. Rather the universe must have begun to exist.

This conclusion has been confirmed by remarkable discoveries in astronomy and astrophysics. The astrophysical evidence indicates that the universe began to exist in a great explosion called the "Big Bang" about 15 billion years ago. Physical space and time were created in that event, as well as all the matter and energy in the universe. Therefore, as Cambridge astronomer Fred Hoyle points out, the Big Bang theory requires the creation of the universe from nothing. This is because, as you go back in time, you reach a point in time at which, in Hoyle's words, the universe was "shrunk down to nothing at all."<sup>\*\*</sup> Thus, what the Big Bang model requires is that the universe began to exist and was created out of nothing.

Now this tends to be very awkward for the atheist. For as Anthony Kenny of Oxford University urges, "A proponent of the Big Bang theory, at least if he is an atheist, must believe that the universe came from nothing and by nothing."<sup>†</sup>

But surely that doesn't make sense! Out of nothing, nothing comes. So why does the universe exist instead of just nothing? Where did it come from? There must have been a cause which brought the universe into being. And from the very nature of the case, this cause must be an uncaused, changeless, timeless, and immaterial being which created the universe. It must be uncaused because there cannot be an infinite regress of causes. It must be timeless and therefore changeless—at least without the universe—because it created time. Because it also created space, it must transcend space as well and therefore be immaterial, not physical.

\* David Hilbert, "On the Infinite," in *Philosophy of Mathematics*, ed. with an Introduction by Paul Benacerraf and Hillary Putnam (Englewood Cliffs, NJ: Prentice-Hall, 1964), 139, 141.

\*\* Fred Hoyle, *Astronomy and Cosmology* (San Francisco: Freeman, 1975), 658.

† Anthony Kenny, *The Five Ways: St. Thomas Aquinas' Proofs of God's Existence* (New York: Schocken Books, 1969), 66.

Moreover, I would argue, it must also be personal. For how else could a timeless cause give rise to a temporal effect like the universe? If the cause were an impersonal set of sufficient conditions, then the cause could never exist without the effect. If the sufficient conditions were timelessly present, then the effect would be timelessly present as well. The only way for the cause to be timeless but for the effect to begin in time is if the cause is a personal agent who freely chooses to create an effect in time without any prior determining conditions. And, thus, we are brought, not merely to the transcendent cause of the universe, but to its personal Creator.

Isn't it incredible that the Big Bang theory thus fits in with what the Christian theist has always believed: that in the beginning God created the universe? Now I put it to you, which do you think makes more sense: that the Christian theist is right or that the universe just popped into being, uncaused, out of nothing? I, at least, have no trouble assessing these alternatives.

2: God makes sense of the complex order in the universe. During the last 30 years, scientists have discovered that the existence of intelligent life depends upon a delicate and complex balance of initial conditions simply given in the Big Bang itself. We now know that life-prohibiting universes are vastly more probable than any life-permitting universe like ours. How much more probable?

Well, the answer is that the chances that the universe should be life-permitting are so infinitesimal as to be incomprehensible and incalculable. For example, Stephen Hawking has estimated that if the rate of the universe's expansion one second after the Big Bang had been smaller by even one part in a hundred thousand million million, the universe would have recollapsed into a hot fireball.\* P. C. W. Davies has calculated that the odds against the initial conditions being suitable for star formation (without which planets could not exist) is one followed by a thousand billion billion zeroes, at least.\*\* [He also] estimates that a change in the strength of gravity or of the weak force by only one part in 10 raised to the 100th power would have prevented a life-permitting universe.<sup>†</sup> There are around 50 such constants and quantities present in the Big Bang which must be fine-tuned in this way if the universe is to permit life. And it's not just each quantity which must be finely tuned; their ratios to each other must also be exquisitely finely tuned. So improbability is multiplied by improbability by improbability until our minds are reeling in incomprehensible numbers.

There is no physical reason why these constants and quantities should possess the values they do. The onetime agnostic physicist P. C. W. Davies comments, "Through my scientific work I have come to believe more and more strongly that the physical universe is put together with an ingenuity

\* Stephen W. Hawking, *A Brief History of Time* (New York: Bantam Books, 1988), 123.

\*\* P. C. W. Davies, *Other Worlds* (London: Dent, 1980), 160–61, 168–69.

† P. C. W. Davies, "The Anthropic Principle," *Particle and Nuclear Physics* 10 (1983): 28.

so astonishing that I cannot accept it merely as a brute fact.”\* Similarly, Fred Hoyle remarks, “A common sense interpretation of the facts suggests that a super-intellect has monkeyed with physics.”\*\* Robert Jastrow, the head of NASA’s Goddard Institute for Space Studies, calls this the most powerful evidence for the existence of God ever to come out of science.<sup>†</sup>

So, once again, the view that Christian theists have always held, that there is an intelligent Designer of the universe, seems to make much more sense than the atheistic interpretation that the universe, when it popped into being, uncaused, out of nothing, just happened to be, by chance, fine-tuned for intelligent life with an incomprehensible precision and delicacy.

3: God makes sense of objective moral values in the world. If God does not exist, then objective moral values do not exist. Many theists and atheists alike concur on this point. For example, the late J. L. Mackie of Oxford University, one of the most influential atheists of our time, admitted: “If . . . there are . . . objective values, they make the existence of a god more probable than it would have been without them. Thus, we have a defensible argument from morality to the existence of God.”<sup>‡</sup> But in order to avoid God’s existence, Mackie therefore denied that objective moral values exist. He wrote, “It is easy to explain this moral sense as a natural product of biological and social evolution.”<sup>§</sup>

Professor Michael Ruse, a philosopher of science at the University of Guelph, agrees. He explains:

Morality is a biological adaptation no less than are hands and feet and teeth. Considered as a rationally justifiable set of claims about an objective something, ethics is illusory. I appreciate that when somebody says, “Love thy neighbor as thyself,” they think they are referring above and beyond themselves. Nevertheless, such reference is truly without foundation. Morality is just an aid to survival and reproduction . . . and any deeper meaning is illusory.<sup>§§</sup>

Friedrich Nietzsche, the great atheist of the last century who proclaimed the death of God, understood that the death of God meant the destruction of all meaning and value in life. I think that Friedrich Nietzsche was right.

But we’ve got to be very careful here. The question here is not: Must we believe in God in order to live moral lives? I’m not claiming that we must. Nor is the question: Can we recognize objective moral values without believing in God? I think we can.

\* Paul Davies, *The Mind of God* (New York: Simon & Schuster, 1992), 169.

\*\* Fred Hoyle, “The Universe: Past and Present Reflections,” *Engineering and Science* 45 (November 1981): 12.

† Robert Jastrow, “The Astronomer and God,” in *The Intellectuals Speak Out About God*, ed. Roy Abraham Varghese (Chicago: Regenry Gateway, 1984), 22.

‡ J. L. Mackie, *The Miracle of Theism* (Oxford: Clarendon Press, 1982), 115–16.

§ Ibid., 117–18.

§§ Michael Ruse, “Evolutionary Theory and Christian Ethics,” in M. Ruse, *The Darwinian Paradigm* (London: Routledge, 1989), 262–69.

Rather the question is: If God does not exist, do objective moral values exist? Like Mackie and Ruse, I just don't see any reason to think that in the absence of God, the morality evolved by homo sapiens is objective. After all, if there is no God, then what's so special about human beings? They're just accidental by-products of nature which have evolved relatively recently on an infinitesimal speck of dust lost somewhere in a hostile and mindless universe and which are doomed to perish individually and collectively in a relatively short time. On the atheistic view, some action, say, rape, may not be socially advantageous, and so in the course of human development has become taboo. But that does absolutely nothing to prove that rape is really wrong. On the atheistic view, there's nothing really wrong with your raping someone. Thus, without God there is no absolute right and wrong which imposes itself on our conscience.

But the problem is that objective moral values do exist, and deep down we all know it. There's no more reason to deny the objective reality of moral values than the objective reality of the physical world. Actions like rape, torture, and child abuse aren't just socially unacceptable behavior; they're moral abominations. Some things, at least, are really wrong. Similarly, love, equality, and self-sacrifice are really good. But if objective values cannot exist without God, and objective values do exist, then it follows logically and inescapably that God exists.

4: God makes sense of the historical facts concerning the life, death and resurrection of Jesus. The historical person, Jesus of Nazareth, was a remarkable individual. New Testament critics have reached something of a consensus that the historical Jesus came on the scene with an unprecedented sense of divine authority, the authority to stand and speak in God's place. That's why the Jewish leadership instigated his crucifixion for the charge of blasphemy. He claimed that in himself the Kingdom of God had come, and as visible demonstrations of this fact, he carried out a ministry of miracle-working and exorcisms. But the supreme confirmation of his claim was his resurrection from the dead. If Jesus did rise from the dead, then it would seem that we have a divine miracle on our hands and, thus, evidence for the existence of God.

Now most people would think that the resurrection of Jesus is just something you believe in by faith or not. But, in fact, there are three established facts, recognized by the majority of New Testament historians today, which I believe support the resurrection of Jesus: the empty tomb; Jesus' postmortem appearances; and the origin of the disciples' belief in his resurrection. Let me say a word about each one of these.

Fact # 1: On the Sunday following his crucifixion, Jesus' tomb was found empty by a group of his women followers. According to Jacob Kremer, an Austrian scholar who has specialized in the study of the resurrection, "By far most scholars hold firmly to the reliability of the Biblical statements about the empty tomb."<sup>\*</sup> According to the New Testament critic, D. H. van Daalen, it is extremely

\* Jacob Kremer, *Die Osterevangelien Geschichten um Geschichte* (Stuttgart: Katholisches Bibelwerk, 1977), 49–50.

difficult to object to the empty tomb on historical grounds; those who deny it do so on the basis of theological or philosophical assumptions.

Fact # 2: On separate occasions different individuals and groups saw appearances of Jesus alive after his death. According to the prominent, skeptical German New Testament critic Gerd Ludemann, “It may be taken as historically certain that . . . the disciples had experiences after Jesus’ death in which Jesus appeared to them as the risen Christ.”<sup>\*</sup> These appearances were witnessed not only by believers, but also by unbelievers, skeptics, and even enemies.

Fact # 3: The original disciples suddenly came to believe in the resurrection of Jesus despite having every predisposition to the contrary. Jews had no belief in a dying, much less a rising, Messiah, and Jewish beliefs about the afterlife precluded anyone’s rising from the dead prior to the end of the world. Luke Johnson, a New Testament scholar at Emory University, muses, “Some sort of powerful, transformative experience is required to generate the sort of movement earliest Christianity was . . .”<sup>\*\*</sup> N. T. Wright, an eminent British scholar, concludes, “That is why, as an historian, I cannot explain the rise of early Christianity unless Jesus rose again, leaving an empty tomb behind him.”<sup>†</sup>

Therefore, it seems to me, the Christian is amply justified in believing that Jesus rose from the dead and was who he claimed to be. But that entails that God exists.

5: God can be immediately known and experienced. This isn’t really an argument for God’s existence; rather it’s the claim that you can know God exists wholly apart from arguments simply by immediately experiencing Him. This was the way people in the Bible knew God, as Professor John Hick explains:

God was known to them as a dynamic will interacting with their own wills, a sheer given reality, as inescapably to be reckoned with as a destructive storm and life-giving sunshine. . . . To them God was not . . . an idea adopted by the mind, but an experiential reality which gave significance to their lives.<sup>‡</sup>

Now if this is so, then there’s a danger that proofs for God could actually distract our attention from God Himself. If you’re sincerely seeking God, then God will make His existence evident to you. The Bible promises, “Draw near to God and He will draw near to you” (James 4:8). We mustn’t so concentrate on the proofs that we fail to hear the inner voice of God speaking to our own heart. For those who listen, God becomes an immediate reality in their lives.

So, in conclusion, we’ve yet to see any arguments to show that God does not exist, and we have seen five reasons to think that God does exist. And, therefore, I think that theism is the more plausible worldview.

\* Gerd Lüdemann, *What Really Happened to Jesus?*, trans. John Bowden (Louisville, KY: Westminster John Knox Press, 1995), 8.

\*\* Luke Timothy Johnson, *The Real Jesus* (San Francisco: Harper San Francisco, 1996), 136.

† N. T. Wright, “The New Unimproved Jesus,” *Christianity Today*, September 13, 1993, 26.

‡ John Hick, “Introduction,” in *The Existence of God*, ed. with an Introduction by John Hick, Problems of Philosophy Series (New York: Macmillan, 1964), 13–14.

**DISCUSSION QUESTIONS**

1. Reconstruct each of Craig's five arguments in standard form.
2. For each of Craig's arguments, what is the main objection that an opponent could raise? What is the best way for Craig to respond to that objection? Is his best response good enough? Why or why not?
3. Craig cites authorities at many points in his speech. Assess his appeals to authorities using the standards in Chapter 15.
4. Is there any strong reason to believe in God that Craig left out? If so, present that reason in argument form as forcefully as you can.

### **SEVEN DEADLY OBJECTIONS TO BELIEF IN THE CHRISTIAN GOD**

by Edwin Curley

Contrary to the impression Dr. Craig's talk may have given you, I am not here tonight to argue for an unqualified atheism. People use the term "God" with many different meanings. Some of them use it to refer to a being whose existence I might be able to accept. If by "God" you mean what I think Spinoza meant by that term—an impersonal system of eternal and immutable laws of nature, which explains everything that happens in the universe, but does not itself require explanation—then I think it quite likely that there is a God. What I'm here to argue against is just the God of Christianity.

I assume that by "God" Christians generally mean an eternal, personal, first cause of the universe, who has infinite perfections—omnipotence, omniscience, perfect goodness, and so on—and who has revealed himself to man in the Christian scriptures. The idea of God as an eternal, personal, first cause of the universe, possessing all perfections, is common not only in Christianity, but also in Judaism and Islam. I take it that the most important thing which distinguishes *Christian* theism from these other forms of theism is the acceptance of the Christian scriptures as the revealed word of God.\* And that's what I shall be arguing against tonight. If the Christian scriptures are supposed to be God's revelation of himself to man, then he is not the perfect being Christians claim he is. . . .

If God is omniscient, if he knows everything, he must have foreknowledge of his creatures' fates. That suggests that there is a fact of the matter about what my fate will be, which is knowable with certainty well in

\* I thought, when I accepted the challenge to debate Dr. Craig, that we had agreed to debate only the existence of the Christian God. This limitation was central to my argumentative strategy here, which relies on the tension between the perfections Christian philosophers and theologians typically ascribe to God and the beliefs they might form about him as a consequence of accepting the Christian scriptures as revealed truth.

Source: Seven Deadly Objections to Belief in the Christian God by Edwin Curley from The Craig-Curley Debate: The Existence of the Christian God. Reprinted by permission.

advance of my birth. It doesn't follow from this that God determined what my fate will be. But it does seem to follow that my fate is not indeterminate. Moreover, if God is omnipotent, and can do anything, or anything it is logically possible to do, then nothing happens except by his will. So if I wind up in Hell, he will not only have known that from eternity, he will also have willed it from eternity.

Predestination is not so widely accepted now as it was when my church was founded in the 16th century. I find many Christians who reject it. And I sympathize with them. Their hearts are in the right place, certainly. I cannot believe that a just and loving God would create beings he knew, and had predetermined, would spend eternity in hell. But Christians can reject predestination only at the cost of ignoring the authority of their scriptures and the implications of their theology. I would not insist that a good Christian must be a fundamentalist, who thinks that everything Scripture says is literally true. But it does seem to me that if someone calls himself a Christian, he ought at least to believe that the Christian scriptures are not seriously misleading about central matters, such as salvation and damnation.

Let us set predestination to one side. Even if the Christian scriptures did not contain that commitment, there would still be a problem about Hell, to which they are committed by an even wider range of texts.\* The philosophical support for Hell, though, seems much weaker than the support for predestination. I see no philosophical reason for believing in an eternal punishment for sinners. Indeed, philosophy is against it. Philosophy teaches that the punishment should be proportionate to the crime. Let's concede, for the sake of argument, that we are all, in some sense, sinners. Which of us, looking into his heart, can honestly say that he has never done anything seriously wrong, at least once in his life? But the doctrine of Hell requires that most of us sinners will suffer eternal torment.\*\* . . .

In the Christian tradition it is normal to baptize infants at an early age because it is believed that they come into the world tainted by the sin of Adam and Eve.<sup>†</sup> This is the doctrine of original sin. I cannot believe in original sin. My granddaughter may be a sinner now, but not when she was in the intensive care unit [for three months after being born prematurely].

\* So far as I can see, the strongest scriptural support for predestination comes from the letters of St. Paul. There seems to be relatively little support for it in the gospels. This might suggest that it was not part of the original teaching of Jesus. But texts supporting the idea of Hell are easy to find in the gospels.

\*\* The most explicit passage I know on the question of the proportion of the saved to the damned is Matthew 22:14, "Many are called, but few are chosen." But if salvation depends on acceptance of Jesus as one's savior, as Christian exclusivists hold, and if the great majority of people have not accepted Jesus as their savior, the texts supporting exclusivism do imply that comparatively few are chosen. I say more about exclusivism below.

<sup>†</sup> This seems to have been the original justification for this practice, though like much else in early Christianity it subsequently became controversial. Some abandoned the practice. Others found other justifications for it. For a good brief account see Alister McGrath, *Christian Theology*, 2nd ed. (Oxford: Blackwell, 1997), 514–18.

Original sin is also less widely accepted now than when my church was founded. I find many Christians who reject original sin. I sympathize with them. Their hearts are in the right place, certainly. But, Christians can reject original sin only at the cost of a substantial reinterpretation of their scriptures and traditions.\*

Consistently with the doctrine of original sin, it is common among Christians to believe that if we are justified, it is by faith in Jesus.\*\* Since we are all sinners, we cannot earn salvation by our works. But we can be forgiven and treated as if we were righteous. The mark of our having been forgiven is that God, by an act of grace, gives us faith.

This doctrine has implications I find appalling. It implies that those among us who lack faith in Jesus have not received grace, have not been forgiven, and will, if we continue in that state, go to Hell. So the doctrine of

\* As Article IX [of the Thirty-nine Articles of Religion of the Episcopal Church] puts it: "Original sin standeth not in the following of Adam (as the Pelagians do vainly talk), but it is the fault and corruption of the Nature of every man that naturally is engendered of the offspring of Adam; whereby man is very far gone from original righteousness, and is of his own nature inclined to evil, so that the flesh lusteth always contrary to the Spirit; and therefore in every person born into this world, it deserveth God's wrath and damnation. And this infection of nature doth remain, yea in them that are regenerated; whereby the lust of the flesh . . . is not subject to the Law of God. And although there is no condemnation for them that believe and are baptized; yet the Apostle doth confess, that concupiscence and lust hath of itself the nature of sin." The principal scriptural authority for this doctrine is found in the letters of St. Paul (the source of so much that is appalling in traditional Christian teaching), notably in Romans 3:9–20, 5:12–21, and 1 Corinthians 15:20–22. I had thought from something Dr. Craig said later in our debate that this doctrine might be rejected in the Wesleyan tradition to which he adheres. But the Articles of Religion of the Methodist Church (available at <http://archives.umc.org/interior.asp?ptid=1&mid=1649>) seem very similar.

\*\* This was certainly the doctrine of my church. See Article XIII: "We are accounted righteous before God, only for the merit of our Lord and Saviour Jesus Christ by Faith, and not for our own works or deservings. Wherefore, that we are justified by Faith only, is a most wholesome Doctrine, and very full of comfort . . ." Scriptural support for this doctrine may be found in Mark 16:16; John 3:16–18, 5:24, 14:6–7; Acts 4:12; Romans 3:9–26, 5:12–21; Ephesians 2:12; I John 2:22–23, 4:3; II John 9. It should, however, be acknowledged that this doctrine seems to be much more common in the works of Paul, and in the fourth gospel, than it is in the synoptic gospels, which tend to suggest that obedience to the commandments, supplemented only by selling all you have and giving it to the poor, is sufficient for salvation. See the story told (with some variations) in Matthew 19:16–22, Mark 10:17–22, Luke 18:18–23. The gospel of John is generally dated late enough that its account of Jesus's teaching might have been influenced by the letters of Paul. (On the dating of Paul's letters and John's gospel, see Raymond Brown's *Introduction to the New Testament*, New Haven, CT: Yale University Press, 1997.) We should concede also that the exceptional passage from Mark comes from the "longer ending of Mark," which is not found in the earliest manuscripts. (See *The New Oxford Annotated Bible*, ed. Bruce Metzger and Roland Murphy, Oxford: Oxford University Press, 1991.) I would not, of course, be thought to suggest that any member of the early Christian Church might have tampered with the text of Mark, to bring it into line with Pauline theology. To do that would be to embrace the kind of radical skepticism for which Bart Ehrman's *Misquoting Jesus* (San Francisco: HarperOne, 2007) has been so widely and justly censured. Still, these textual data are puzzling. A Christian who took his scriptures seriously might think they require some explanation.

justification by faith, which has strong support in the Christian scriptures, leads inevitably to exclusivism, to the idea that all who reject Christian doctrine must be damned, no matter how good they may be in other respects.

If God chose the beneficiaries of his grace on the ground of some distinctive merit they possessed, this might not be unfair to those he didn't choose, whom we would presume to lack that merit. But that would be contrary to the idea of grace, which implies a free gift, not something given to someone who deserves it on account of merit.

So usually it is held that God has no reason for choosing some and not others. He acts quite arbitrarily. It's a hard and ugly doctrine, this doctrine of grace. Of course if you have already accepted Hell and original sin, you may find this doctrine "full of comfort," as my church put it. You may be grateful for having this chance at salvation, even if it does seem to be a lottery in which the odds are not on your side. And if you think you have faith, then you may also think you have won the lottery. Still even the faithful may not be able to entirely set aside thoughts about the unlucky losers. Sometimes they love non-believers and then the doctrine of salvation by faith can cause them great grief.

So far my objections have been mainly theological; they are objections to teachings whose basis is primarily scriptural rather than philosophical. The main exception to that generalization is the doctrine of predestination, which has philosophical grounds as well as scriptural grounds. I know many Christians here tonight will not feel that their understanding of Christianity requires them to accept all these doctrines, either because they do not think their scriptures clearly require these views, or because they do not regard the Christian scriptures as absolutely authoritative in determining their beliefs and conduct. I've said I think those Christians who adopt a freer attitude toward scripture and do not feel that their acceptance of Christianity commits them to predestination, or Hell, or original sin, or justification by faith, or exclusivism—those Christians have their hearts in the right place. But I also think their feet may be planted on the slippery slope to heresy, and that more conservative Christians, who would accord greater authority to scripture, have a clearer right to consider themselves Christians. How much of traditional Christianity can you reject and still be entitled to call yourself a Christian? This was a question I found extremely difficult as I moved gradually from Christianity to apostasy.

Let's turn now to objections not so scripturally based. It is common among Christians to believe that God is a personal being, who created the universe, and who is omnipotent, omniscient, and perfectly good. Indeed, it is commonly said that God must possess all perfections.

Yet we observe that the world this perfect being created has many imperfections: there is much joy in the world; but there is also much suffering, much of it apparently undeserved; and there is sin. We call these things evil. How can they exist in a world which owes its origin to a God with the attributes Christians believe their God to possess?

The usual response now is to say that, though God could have created a world without evil, it was better for him to have created the world he did, in spite of the evils it contains. The occurrence of those evils was necessary for goods which are even greater. If God had so created the world that it contained no evil at all, that world would have been less good, all things considered, than it is even with all the evil it contains. This is called the greater goods defense.

The Christian may say: we humans rightly do many things we expect to cause avoidable harm. We build a bridge from San Francisco to Marin County, knowing that in the construction some workmen will fall into the water and die. We could avoid their deaths by not building the bridge. But the bridge is a great good. Given our human limitations, we cannot build it without some people dying a result. So we build it and accept their deaths as part of the cost of bridging those waters. And God's permission of evil may also be justified by the greater goods it leads to.

An omnipotent being, of course, does not face all the hard choices we do. If he wants a bridge across those waters, he need only say, "Let there be a bridge." And there will be. This fact that God is supposed to be omnipotent puts constraints on the kinds of good he might be able to use to justify his acceptance of the evils he allows. What kind of good could be so intimately connected with evil that even an omnipotent being would have to accept the evil, as the price of realizing that good? And what good could be so great that it would justify such a being's accepting the amount of evil there is in the world as the price of attaining that good?

The usual answer is: freedom. There must be freedom, if there is to be moral goodness. And the price of giving humans freedom is that sometimes they will misuse it. Even an omnipotent being can't cause a person to freely do good. If he caused one of his creatures to behave well, that person would no longer be acting freely. His "acts" would no longer be *his* acts, and he would not deserve moral credit for them. Freedom, with the moral goodness which sometimes results from it, is a good sufficiently great that it makes the evils which also result worth accepting. This is what is called the free will defense.

There is a problem, of course, about appealing to human freedom to solve the problem of evil when you also believe in divine foreknowledge and predestination. This is a problem of long standing, which many philosophers have wrestled with. No solution has gained general acceptance. If Dr. Craig accepts the doctrines of predestination and divine foreknowledge and also appeals to human freedom to solve the problem of evil, he will have worked out a way of explaining how these things are consistent, and I will listen with interest to that explanation.

In the meantime, though, there are other problems about the appeal to freedom. There are evils whose occurrence has no discernible connection with freedom. Theologians call them natural evils, meaning by that such things as earthquakes, floods, tornadoes, diseases, and so on. If a deer dies

in a forest fire, suffering horribly as it does so, that is an evil. It is not only human suffering we must take into account, when we are weighing good against evil in this world.

Now, if you accept anything like the theory of evolution,\* you will believe there were other animals on this planet long before humans appeared on the scene. Many of them must have suffered horribly as their species became extinct. None of that suffering can be justified as a necessary consequence of permitting humans freedom. We weren't around then. So none of it seems beyond the power of omnipotence to prevent without the loss of that good.

Another objection: The greater goods defense can easily lead to a kind of cost-benefit analysis which is deeply repugnant to our moral sense. Consider the kind of case which troubled Ivan in Dostoevsky's great novel, *The Brothers Karamazov*.\*\* A little girl is treated quite brutally by her parents, who beat her because she has done something which made them angry. Perhaps she wets the bed repeatedly, and they think she ought to be old enough to control her bladder. Or perhaps the father is an alcoholic who abuses his daughter sexually. *The Brothers Karamazov* is fiction, but to hear about real cases like this, you need only listen regularly to the 11 o'clock news.

The free will defense seems to say, in cases of this kind: well, it's all very unfortunate, of course, but this is the price we must pay for having freedom. For the father to have the opportunity to display moral goodness, God must give him the opportunity to choose evil. You can't have the one opportunity without the other. And the father's having the opportunity to display moral goodness is such a great good that it outweighs the fact that he chooses evil.

But notice who gets the good here. It's the father. And notice who suffers the evil. It's the little girl. Let us grant, for the sake of argument, that the benefit outweighs the cost. Freedom is a very great good. Still it makes some difference who pays the cost. Freedom may be a great good, even a good so great that it would outweigh really horrendous suffering. But justice requires some attention, not only to the net amount of good, after you have subtracted the evil, but also to the way the goods and evils are distributed. Some distributions just aren't fair.<sup>†</sup>

\* You don't, of course, have to accept the full Darwinian explanation of evolution for this to be a problem, so long as you accept what the evidence of geology and paleontology seems to make as certain as anything in science can be: that the earth has been around for a very long time, and that many, many different species of animals flourished and then became extinct before man appeared on the scene.

\*\* See Part II, Book V: Pro and Contra, Chapter 4: Rebellion. The text is available online at: <http://www.online-literature.com/booksearch.php>.

† I first became aware of the importance of this by reading Michael Tooley's "The Argument from Evil," in *Philosophical Perspectives*, 5—*Philosophy of Religion*, 1991, ed. James E. Tomberlin (Atascadero, CA: Ridgeview Publishing, 1991), 89–134. But I think acknowledgment of this point underlies Marilyn Adams's rejection of the free will defense in her *Horrendous Evils and the Goodness of God* (Ithaca, NY: Cornell University Press, 2000).

The mention of Ivan Karamazov brings me to my final objection. Ivan claims that if God does not exist, everything is permissible. Dr. Craig believes the same thing. Dostoevsky, speaking through Ivan, may have stated the problem of evil as powerfully as any atheist; but he was himself a Christian, who believed that God must exist if we are to make sense of morality.

I think the opposite is true. I think Christian belief makes morality, as we normally think of it, unintelligible. Consider the story of Abraham and Isaac. One day God put Abraham to the test. He said to Abraham: "Take your only son, Isaac, whom you love, and go to the land of Moriah, and offer him there as a burnt offering."<sup>\*</sup> God gives no reason for this horrifying command. And Abraham asks none. He simply sets out to obey the command. And he nearly does obey. He has the knife raised to kill his son, when God sends down an angel to stay his hand. God then says he is satisfied with Abraham: "Now I know that you fear God, since you have not withheld your son, your only son, from me" (Gen. 22:12). In the end God does not actually require the sacrifice. But he does require that Abraham demonstrate his willingness to carry out the sacrifice.

What's the moral of this story? I suggest it's this: as God's creatures, our highest loyalty must be to God, even if this requires the sacrifice of our deepest human loyalties; God is our Creator, our Lord, and we owe him absolute obedience, no matter what he commands. And he might command anything. There are no constraints on his will; so we might be required to do anything. There is no predicting what he might require; and there is nothing to say that his commands will not change from one moment to the next. At the beginning of the story, God commands Abraham to kill Isaac; in the middle he commands Abraham not to kill Isaac.

If there is a God who is liable to command anything; and if our highest loyalty must be to this God, there is no act save disobedience to God which we can safely say is out of bounds, no act of a kind which simply must not be done, even rape, to use Dr. Craig's example. If this God exists, and we must obey him unconditionally, then anything whatever might turn out to be permissible. This view is destructive of morality as we normally think of it.

So there you have my opening argument. I have offered seven objections, seven deadly objections, I would say, to Christian theism: it is committed to predestination, to Hell, to original sin, to justification by faith, and to exclusivism; it has no good solution to the problem of evil; and it is destructive of morality as we understand it. These are only some of the objections which make it impossible for me to believe in the Christian God. But they are enough to make me wonder how anyone who has thought seriously about the Christian faith can accept it.

\*Genesis 22:2, quoted in the New Revised Standard Version, as given in *The New Oxford Annotated Bible*.

**DISCUSSION QUESTIONS**

1. Explain and evaluate each of Curley's "seven deadly objections" to belief in a Christian God. How could a traditional Christian best respond to each objection? Is that response adequate? Why or why not?
2. Some of Curley's arguments depend on specific doctrines of traditional Christianity. Does this make his arguments weaker? Why or why not?
3. Is there any better reason to deny the existence of God that Curley overlooked? If so, present it as forcefully as you can.



## PHILOSOPHICAL REASONING

*It is not easy to explain the character of philosophical reasoning. One source of confusion is that philosophers reason in very different ways. They use diverse methods to explore diverse topics. In addition, philosophers reason about reasoning, and they often disagree about what counts as proper reasoning. As a result, the nature of philosophical reasoning is itself a philosophical problem. This chapter, then, cannot cover all of philosophical reasoning. Instead, we will focus on one example—free will—that brings together many features that are central to much philosophical reasoning.*

We can acquire some sense of philosophical reasoning by contrasting it with reasoning in other areas. Reasoning in religion usually takes for granted that certain scriptures and leaders have special authority. Premises in religious reasoning can then cite what is said in those texts or by those people. Similarly, reasoning in law usually assumes that certain institutions (such as legislatures and courts) as well as texts (such as constitutions, statutes, and opinions in precedents) have special authority and can serve as the basis for legal reasoning. Moreover, this book has often emphasized that, in everyday discussions, much is taken for granted and left unsaid. In general, there is no need to state points that are already a matter of agreement.

In contrast, philosophers usually try to make underlying assumptions explicit and then subject them to critical examination. All premises and presuppositions should be brought into the light of day, so that we can ask what exactly they mean and whether we ought to accept them. No authority may be taken for granted without support. No assumption is sacrosanct, even the assumption that no assumption is sacrosanct, according to most philosophers.

Nonetheless, even for the philosopher, something must trigger an interest in a particular underlying assumption. This usually arises when the advance of knowledge creates fundamental conflicts within the system of hitherto accepted assumptions or so-called common sense. Thus, much that counts as modern philosophy is an attempt to come to terms with the relationship between modern science and the traditional conception of humankind's place in the universe. Conflicts between science and common sense feed

philosophy. This chapter explores one such conflict between our traditional sense of personal responsibility and recent advances in physics, psychology, sociology, and neuroscience.

Modern sciences have attempted to understand human behavior by looking for causes of what we do. First came physics. Our bodies are made of atoms, and physicists discovered laws that govern the movements of those atoms, so our bodily movements seem to be determined by laws of physics. Then psychologists discovered that our actions often result from unconscious influences, and sociologists found many ways in which we are affected by social environments. Most recently, neuroscientists have traced our actions to chemical and electrical mechanisms in our brains. Together, all of this science might seem to suggest a picture of people as puppets controlled by outside influences.

The problem is that we often see ourselves as free to choose among available options and free to act as we choose. We also hold most people responsible for choosing and acting as they do. What is not clear is how we can be free and responsible if we really are just puppets controlled by outside influences. Thus, science seems to conflict with common sense in this area.

These issues matter deeply because they influence our self-perception as well as our practice. Without free will, are we really any different from lower animals, like dogs? If we are puppets of our circumstances, do our lives have any more meaning than the lives of dogs? And if we are not free or responsible, can it really be fair to punish rapists and murderers? Can it make sense to love our friends, feel guilty when we misbehave, or get angry at people who hurt us? Our lives and institutions seem unjustified or even incoherent if we cannot resolve this conflict between science and common sense.

In order to figure out how to solve these problems, we need to clarify exactly where the problems lie. There are at least three distinct conflicts in this general area. The first problem involves *determinism*, which is the claim that all of our choices and actions are fully caused and determined by laws of nature in conjunction with the initial conditions before we were born (say, in 1900). Determinism seems to imply that we *cannot* act in any way other than the way we actually do act. This lack of alternative possibility seems to suggest that we are neither free nor responsible.

Philosophers have proposed several solutions to this problem. The simplest is to reject the commonsense assumptions that we are free and responsible. That view follows from the premises that we are determined and that determinism is incompatible with freedom and responsibility. This solution is called *hard determinism*. Philosophers who want to hold on to freedom and responsibility need to reject hard determinism, and they have several options.

Some philosophers try to save freedom and responsibility by denying determinism. Quantum mechanics famously rejects determinism in physics. However, quantum mechanics replaces determinism with randomness, and that kind of indeterminism does not seem adequate to secure freedom or

responsibility. If my actions are produced by random quantum fluctuations of which I have no awareness or control, then it is hard to see how that kind of indeterminacy could make me free or responsible for what those random events cause. For this reason, some philosophers claim that both determinism and also this kind of indeterminism rule out freedom and responsibility. If they go on to deny freedom and responsibility, they can remain neutral about whether the universe is deterministic, and then they can end up in the position that Derk Pereboom calls *hard incompatibilism* (in the reading below).

Other philosophers, in contrast, claim that people can choose and act without being caused or at least determined by any outside influences and also without those choices being random, because they are based on reasons. In this view, our choices are not caused by anything other than ourselves as agents, and then our choices cause our actions. Moreover, we make those choices (at least sometimes) for reasons that are not reducible to causes. This view is usually called *libertarianism* (though it should not be confused with the political view of the same name).

Both hard determinists and libertarians assume that determinism is incompatible with free will or free action. That shared assumption is questioned by *compatibilism*, which is the position that we can be both fully determined and also free and responsible. Compatibilists claim that there is no real conflict between science and common sense. You can have both, in their view. Some particular versions of compatibilism are discussed by Susan Wolf and by Derk Pereboom in their essays reprinted in this chapter. When compatibilism is combined with determinism, the conjunction is then called *soft determinism*.

We end up with three main options: hard determinism (or hard incompatibilism), libertarianism, and compatibilism (or soft determinism). Philosophers have recently created variations on each of these options, such as *semi-compatibilism*, which claims that determinism is compatible with responsibility but not with freedom. Still, adding more options does not resolve these debates. Resolution requires argument.

Although the problem of determinism has been central to traditional discussions of free will and responsibility, recent philosophers have distinguished the problem of determinism from other challenges raised by science. One of these new challenges is *bypassing* or *mechanism*. The idea is that freedom and responsibility are challenged not by determinism but by the purely mechanical or mind-free nature of causation in science. If physics or neuroscience can explain our bodily movements by citing atoms and neurons without ever mentioning any mental state, such as a desire or choice, then the causal chain that leads to our actions seems to bypass our minds. Our mental states, including our desires and choices, seem unnecessary, impotent, or *epiphenomenal* (in the sense that physical events cause mental events, but mental events never cause physical events). The problem is that it is hard to see how we can be free or responsible for how our bodies move

if our choices play no causal role in bringing about those bodily movements. We seem no more free than a person having an epileptic seizure, whose bodily movements do not result from will or choice. This second challenge arises from the tendency of scientific explanations, at least in physics and neuroscience, to work without mentioning mental states at all.

A third problem for free will and responsibility concerns the *sources* of our mental states and actions. Psychologists and sociologists emphasize that the physical, social, and cultural environment in which we make our choices influences what we choose. They sometimes go further and suggest that our choices and actions are determined by our environment, such that anybody else would have done the same thing in the same circumstances. If so, then people who act badly are merely unlucky to have encountered the circumstances that caused them to act badly. Their bad actions do not show anything bad about them or their characters if anybody else in their circumstances would have done the same bad acts. Moreover, we do not choose and are not responsible for the environment in which we live. That makes it hard to understand how we could be free or responsible for choices and actions that result from the unchosen environment. Of course, opponents respond that our environments or circumstances do not totally determine what we choose or do, so at least part of the source of our actions lies in our selves—or perhaps in some special part of our selves called the *deep self*. Then these philosophers argue that we are responsible when our actions are caused by a deep self with certain special properties, such as rationality or responsiveness to reasons or to truth and goodness.

This kind of view is developed by Susan Wolf in our readings. Wolf describes and criticizes some views of the “deep self” and then argues for her own variation, which is supposed to explain how and why people can be free and responsible despite all of science. Derk Pereboom then responds that any deep self can be manipulated and caused in ways that are incompatible with real freedom and responsibility. As you read through these selections, it will be useful to keep in mind all three of the challenges to free will and responsibility: the problems of determinism, of bypassing, and of sourcehood. Ask yourself which issue is being addressed at each point in the readings.

#### DISCUSSION QUESTIONS

1. Which is the best solution to the problem of determinism: hard determinism, libertarianism, or compatibilism? Why?
2. If a neuroscientist fully explained a person’s bodily movements by citing only electrical and chemical events in that person’s brain (along with energy and matter entering that person’s body from outside), would that explanation convince you that the person’s mental states did not cause the bodily movements (so the mental states are “bypassed”)? Would it convince you that the person is not free and responsible? Why or why not?

## SANITY AND THE METAPHYSICS OF RESPONSIBILITY<sup>1</sup>

by Susan Wolf

Philosophers who study the problems of free will and responsibility have an easier time than most in meeting challenges about the relevance of their work to ordinary, practical concerns. Indeed, philosophers who study these problems are rarely faced with such challenges at all, since questions concerning the conditions of responsibility come up so obviously and so frequently in everyday life. Under scrutiny, however, one might question whether the connections between philosophical and nonphilosophical concerns in this area are real.

In everyday contexts, when lawyers, judges, parents, and others are concerned with issues of responsibility, they know, or think they know, what in general the conditions of responsibility are. Their questions are questions of application: Does this or that particular person meet this or that particular condition? Is this person mature enough, or informed enough, or sane enough to be responsible? Was he or she acting under posthypnotic suggestion or under the influence of a mind-impairing drug? It is assumed, in these contexts, that normal, fully developed adult human beings are responsible beings. The questions have to do with whether a given individual falls within the normal range.

By contrast, philosophers tend to be uncertain about the general conditions of responsibility, and they care less about dividing the responsible from the nonresponsible agents than about determining whether, and if so why, any of us are ever responsible for anything at all.

In the classroom, we might argue that the philosophical concerns grow out of the nonphilosophical ones, that they take off where the nonphilosophical questions stop. In this way, we might convince our students that even if they are not plagued by the philosophical worries, they ought to be. If they worry about whether a person is mature enough, informed enough, and sane enough to be responsible, then they should worry about whether that person is metaphysically free enough, too.

The argument I make here, however, goes in the opposite direction. My aim is not to convince people who are interested in the apparently nonphilosophical conditions of responsibility that they should go on to worry about the philosophical conditions as well, but rather to urge those who already worry about the philosophical problems not to leave the more mundane, prephilosophical problems behind. In particular, I suggest that the mundane recognition that *sanity* is a condition of responsibility has more to do with the murky and apparently metaphysical problems which surround the issue of responsibility than at first meets the eye. Once the significance of the condition of sanity is fully appreciated, at least some of the apparently insuperable metaphysical aspects of the problem of responsibility will dissolve.

My strategy is to examine a recent trend in philosophical discussions of responsibility, a trend that tries, but I think ultimately fails, to give an acceptable analysis of the conditions of responsibility. It fails due to what at first appear to be deep and irresolvable metaphysical problems. It is here that I suggest that the condition of sanity comes to the rescue. What at first appears to be an impossible requirement for responsibility—the requirement that the responsible agent have created her- or himself—turns out to be the vastly more mundane and noncontroversial requirement that the responsible agent must, in a fairly standard sense, be sane.

#### FRANKFURT, WATSON, AND TAYLOR

The trend I have in mind is exemplified by the writings of Harry Frankfurt, Gary Watson, and Charles Taylor. I will briefly discuss each of their separate proposals, and then offer a composite view that, while lacking the subtlety of any of the separate accounts, will highlight some important insights and some important blind spots they share.

In his seminal article “Freedom of the Will and the Concept of a Person,”<sup>1</sup> Harry Frankfurt notes a distinction between freedom of action and freedom of the will. A person has freedom of action, he points out, if she (or he) has the freedom to do whatever she wills to do—the freedom to walk or sit, to vote liberal or conservative, to publish a book or open a store, in accordance with her strongest desires. Even a person who has freedom of action may fail to be responsible for her actions, however, if the wants or desires she has the freedom to convert into action are themselves not subject to her control. Thus, the person who acts under posthypnotic suggestion, the victim of brainwashing, and the kleptomaniac might all possess freedom of action. In the standard contexts in which these examples are raised, it is assumed that none of the individuals is locked up or bound. Rather, these individuals are understood to act on what, at one level at least, must be called *their own desires*. Their exemption from responsibility stems from the fact that their own desires (or at least the ones governing their actions) are not up to them. These cases may be described in Frankfurt’s terms as cases of people who possess freedom of action, but who fail to be responsible agents because they lack freedom of the will.

Philosophical problems about the conditions of responsibility naturally focus on an analysis of this latter kind of freedom: What is freedom of the will, and under what conditions can we reasonably be thought to possess it? Frankfurt’s proposal is to understand freedom of the will by analogy to freedom of action. As freedom of action is the freedom to do whatever one wills to do, freedom of the will is the freedom to will whatever one wants to will. To make this point clearer, Frankfurt introduces a distinction between

<sup>1</sup> Harry Frankfurt, “Freedom of the Will and the Concept of a Person,” *Journal of Philosophy* LXVIII (1971), 5–20.

first-order and second-order desires. First-order desires are desires to do or to have various things; second-order desires are desires about what desires to have or what desires to make effective in action. In order for an agent to have both freedom of action and freedom of the will, that agent must be capable of governing his or her actions by first-order desires *and* capable of governing his or her first-order desires by second-order desires.

Gary Watson's view of free agency<sup>2</sup>—free and responsible agency, that is—is similar to Frankfurt's in holding that an agent is responsible for an action only if the desires expressed by that action are of a particular kind. While Frankfurt identifies the right kind of desires as desires that are supported by second-order desires, however, Watson draws a distinction between "mere" desires, so to speak, and desires that are *values*. According to Watson, the difference between free action and unfree action cannot be analyzed by reference to the logical form of the desires from which these various actions arise, but rather must relate to a difference in the quality of their source. Whereas some of my desires are just appetites or conditioned responses I find myself "stuck with," others are expressions of judgments on my part that the objects I desire are good. Insofar as my actions can be governed by the latter type of desire—governed, that is, by my values or valuational system—they are actions that I perform freely and for which I am responsible.

Frankfurt's and Watson's accounts may be understood as alternate developments of the intuition that in order to be responsible for one's actions, one must be responsible for the self that performs these actions. Charles Taylor, in an article entitled "Responsibility for Self,"<sup>3</sup> is concerned with the same intuition. Although Taylor does not describe his view in terms of different levels or types of desire, his view is related, for he claims that our freedom and responsibility depends on our ability to reflect on, criticize, and revise our selves. Like Frankfurt and Watson, Taylor seems to believe that if the characters from which our actions flowed were simply and permanently *given* to us, implanted by heredity, environment, or God, then we would be mere vehicles through which the causal forces of the world traveled, no more responsible than dumb animals or young children or machines. But like the others, he points out that, for most of us, our characters and desires are not so brutally implanted—or, at any rate, if they are, they are subject to revision by our own reflecting, valuing, or second-order desiring selves. We human beings—and as far as we know, only we human beings—have the ability to step back from ourselves and decide whether we are the selves we want to be. Because of this, these philosophers think, we are responsible for our selves and for the actions that we produce.

Although there are subtle and interesting differences among the accounts of Frankfurt, Watson, and Taylor, my concern is with features of

<sup>2</sup> Gary Watson, "Free Agency," *Journal of Philosophy* LXXII (1975), 205–20.

<sup>3</sup> Charles Taylor, "Responsibility for Self," in A. E. Rorty, ed., *The Identities of Persons* (Berkeley: University of California Press, 1976), pp. 281–99.

their views that are common to them all. All share the idea that responsible agency involves something more than intentional agency. All agree that if we are responsible agents, it is not just because our actions are within the control of our wills, but because, in addition, our wills are not just psychological states *in us*, but expressions of characters that come *from us*, or that at any rate are acknowledged and affirmed *by us*. For Frankfurt, this means that our wills must be ruled by our second-order desires; for Watson, that our wills must be governable by our system of values; for Taylor, that our wills must issue from selves that are subject to self-assessment and redefinition in terms of a vocabulary of worth. In one way or another, all these philosophers seem to be saying that the key to responsibility lies in the fact that responsible agents are those for whom it is not just the case that their actions are within the control of their wills, but also the case that their wills are within the control of their *selves* in some deeper sense. Because, at one level, the differences among Frankfurt, Watson, and Taylor may be understood as differences in the analysis or interpretation of what it is for an action to be under the control of this deeper self, we may speak of their separate positions as variations of one basic view about responsibility: the *deep-self view*.

#### THE DEEP-SELF VIEW

Much more must be said about the notion of a deep self before a fully satisfactory account of this view can be given. Providing a careful, detailed analysis of that notion poses an interesting, important, and difficult task in its own right. The degree of understanding achieved by abstraction from the views of Frankfurt, Watson, and Taylor, however, should be sufficient to allow us to recognize some important virtues as well as some important drawbacks of the deep-self view.

One virtue is that this view explains a good portion of our pretheoretical intuitions about responsibility. It explains why kleptomaniacs, victims of brainwashing, and people acting under posthypnotic suggestion may not be responsible for their actions, although most of us typically are. In the cases of people in these special categories, the connection between the agents' deep selves and their wills is dramatically severed—their wills are governed not by their deep selves, but by forces external to and independent from them. A different intuition is that we adult human beings can be responsible for our actions in a way that dumb animals, infants, and machines cannot. Here the explanation is not in terms of a split between these beings' deep selves and their wills; rather, the point is that these beings *lack* deep selves altogether. Kleptomaniacs and victims of hypnosis exemplify individuals whose selves are *alienated* from their actions; lower animals and machines, on the other hand, do not have the sorts of selves from which actions *can* be alienated, and so they do not have the sort of selves from which, in the happier cases, actions can responsibly flow.

At a more theoretical level, the deep-self view has another virtue: It responds to at least one way in which the fear of determinism presents itself.

A naive reaction to the idea that everything we do is completely determined by a causal chain that extends backward beyond the times of our births involves thinking that in that case we would have no control over our behavior whatsoever. If everything is determined, it is thought, then what happens happens, whether we want it to or not. A common, and proper, response to this concern points out that determinism does not deny the causal efficacy an agent's desires might have on his or her behavior. On the contrary, determinism in its more plausible forms tends to affirm this connection, merely adding that as one's behavior is determined by one's desires, so one's desires are determined by something else.<sup>4</sup>

Those who were initially worried that determinism implied fatalism, however, are apt to find their fears merely transformed rather than erased. If our desires are governed by something else, they might say, they are not *really* ours after all—or, at any rate, they are ours in only a superficial sense.

The deep-self view offers an answer to this transformed fear of determinism, for it allows us to distinguish cases in which desires are determined by forces foreign to oneself from desires which are determined *by* one's self—by one's “real,” or second-order desiring, or valuing, or deep self, that is. Admittedly, there are cases, like that of the kleptomaniac or the victim of hypnosis, in which the agent acts on desires that “belong to” him or her in only a superficial sense. But the proponent of the deep-self view will point out that even if determinism is true, ordinary adult human action can be distinguished from this. Determinism implies that the desires which govern our actions are in turn governed by something else, but that something else will, in the fortunate cases, be our own deeper selves.

This account of responsibility thus offers a response to our fear of determinism; but it is a response with which many will remain unsatisfied. Even if my actions are governed by my desires and my desires are governed by my own deeper self, there remains the question: Who, or what, is responsible for this deeper self? The response above seems only to have pushed the problem further back.

Admittedly, some versions of the deep-self view, including Frankfurt's and Taylor's, seem to anticipate this question by providing a place for the ideal that an agent's deep self may be governed by a still deeper self. Thus, for Frankfurt, second-order desires may themselves be governed by third-order desires, third-order desires by fourth-order desires, and so on. Also, Taylor points out that, as we can reflect on and evaluate our prereflective selves, so we can reflect on and evaluate the selves who are doing the first

<sup>4</sup> See, e.g., David Hume, *A Treatise of Human Nature* (Oxford: Oxford University Press, 1967), pp. 399–406, and R. E. Hobart, “Free Will as Involving Determination and Inconceivable Without It,” *Mind* 43 (1934).

reflecting and evaluating, and so on. However, this capacity to recursively create endless levels of depth ultimately misses the criticism's point.

First of all, even if there is no *logical* limit to the number of levels of reflection or depth a person may have, there is certainly a psychological limit—it is virtually impossible imaginatively to conceive a fourth-, much less an eighth-order, desire. More important, no matter how many levels of self we posit, there will still, in any individual case, be a last level—a deepest self about whom the question “What governs it?” will arise, as problematic as ever. If determinism is true, it implies that even if my actions are governed by my desires, and my desires are governed by my deepest self, my deepest self will still be governed by something that must, logically, be external to myself altogether. Though I can step back from the values my parents and teachers have given me and ask whether these are the values I really want, the “I” that steps back will itself be a product of the parents and teachers I am questioning.

The problem seems even worse when one sees that one fares no better if determinism is false. For if my deepest self is not determined by something external to myself, it will still not be determined by *me*. Whether I am a product of carefully controlled forces or a result of random mutations, whether there is a complete explanation of my origin or no explanation at all, I am not, in any case, responsible for my existence; I am not in control of my deepest self.

Thus, though the claim that an agent is responsible for only those actions that are within the control of his or her deep self correctly identifies a necessary condition for responsibility—a condition that separates the hypnotized and the brainwashed, the immature and the lower animals from ourselves, for example—it fails to provide a sufficient condition of responsibility that puts all fears of determinism to rest. For one of the fears invoked by the thought of determinism seems to be connected to its implication that we are but intermediate links in a causal chain, rather than ultimate, self-initiating sources of movement and change. From the point of view of one who has this fear, the deep-self view seems merely to add loops to the chain, complicating the picture but not really improving it. From the point of view of one who has this fear, responsibility seems to require being a prime mover unmoved, whose deepest self is itself neither random *nor* externally determined, but is rather determined *by itself*—who is, in other words, self-created.

At this point, however, proponents of the deep-self view may wonder whether this fear is legitimate. For although people evidently can be brought to the point where they feel that responsible agency requires them to be ultimate sources of power, to the point where it seems that nothing short of self-creation will do, a return to the internal standpoint of the agent whose responsibility is in question makes it hard to see what good this metaphysical status is supposed to provide or what evil its absence is supposed to impose.

From the external standpoint, which discussions of determinism and indeterminism encourage us to take up, it may appear that a special metaphysical status is required to distinguish us significantly from other members of the natural world. But proponents of the deep-self view will suggest this is an illusion that a return to the internal standpoint should dispel. The possession of a deep self that is effective in governing one's actions is a sufficient distinction, they will say. For while other members of the natural world are not in control of the selves that they are, we, possessors of effective deep selves, are in control. We can reflect on what sorts of beings we are, and on what sorts of marks we make on the world. We can change what we don't like about ourselves, and keep what we do. Admittedly, we do not create ourselves from nothing. But as long as we can revise ourselves, they will suggest, it is hard to find reason to complain. Harry Frankfurt writes that a person who is free to do what he wants to do and also free to want what he wants to want has "all the freedom it is possible to desire or to conceive."<sup>5</sup> This suggests a rhetorical question: If you are free to control your actions by your desires, and free to control your desires by your deeper desires, and free to control those desires by still deeper desires, what further kind of freedom can you want?

#### THE CONDITION OF SANITY

Unfortunately, there is a further kind of freedom we can want, which it is reasonable to think necessary for responsible agency. The deep-self view fails to be convincing when it is offered as a complete account of the conditions of responsibility. To see why, it will be helpful to consider another example of an agent whose responsibility is in question.

JoJo is the favorite son of Jo the First, an evil and sadistic dictator of a small, undeveloped country. Because of his father's special feelings for the boy, JoJo is given a special education and is allowed to accompany his father and observe his daily routine. In light of this treatment, it is not surprising that little JoJo takes his father as a role model and develops values very much like Dad's. As an adult, he does many of the same sorts of things his father did, including sending people to prison or to death or to torture chambers on the basis of whim. He is not *coerced* to do these things, he acts according to his own desires. Moreover, these are desires he wholly *wants* to have. When he steps back and asks, "Do I really want to be this sort of person?" his answer is resoundingly "Yes," for this way of life expresses a crazy sort of power that forms part of his deepest ideal.

In light of JoJo's heritage and upbringing—both of which he was powerless to control—it is dubious at best that he should be regarded as responsible for what he does. It is unclear whether anyone with a childhood such as his could have developed into anything but the twisted and perverse sort

<sup>5</sup> Frankfurt, p. 16.

of person that he has become. However, note that JoJo is someone whose actions are controlled by his desires and whose desires are the desires he wants to have: That is, his actions are governed by desires that are governed by and expressive of his deepest self.

The Frankfurt–Watson–Taylor strategy that allowed us to differentiate our normal selves from the victims of hypnosis and brainwashing will not allow us to differentiate ourselves from the son of Jo the First. In the case of these earlier victims, we were able to say that although the actions of these individuals were, at one level, in control of the individuals themselves, these individuals themselves, *qua* agents, were not the selves they more deeply wanted to be. In this respect, these people were unlike our happily more integrated selves. However, we cannot say of JoJo that his self, *qua* agent, is not the self he wants it to be. It *is* the self he wants it to be. From the inside, he feels as integrated, free, and responsible as we do.

Our judgment that JoJo is not a responsible agent is one that we can make only from the outside—from reflecting on the fact, it seems, that his deepest self is not up to him. Looked at from the outside, however, our situation seems no different from his—for in the last analysis, it is not up to any of us to have the deepest selves we do. Once more, the problem seems metaphysical—and not just metaphysical, but insuperable. For, as I mentioned before, the problem is independent of the truth of determinism. Whether we are determined or undetermined, we cannot have created our deepest selves. Literal self-creation is not just empirically, but logically impossible.

If JoJo is not responsible because his deepest self is not up to him, then we are not responsible either. Indeed, in that case responsibility would be impossible for anyone to achieve. But I believe the appearance that literal self-creation is required for freedom and responsibility is itself mistaken.

The deep-self view was right in pointing out that freedom and responsibility requires us to have certain distinctive types of control over our behavior and our selves. Specifically, our actions need to be under the control of our selves, and our (superficial) selves need to be under the control of our deep selves. Having seen that these types of control are not enough to guarantee us the status of responsible agents, we are tempted to go on to suppose that we must have yet another kind of control to assure us that even our deepest selves are somehow up to us. But not all the things necessary for freedom and responsibility must be types of power and control. We may need simply to *be* a certain way, even though it is not within our power to determine whether we are that way or not.

Indeed, it becomes obvious that at least one condition of responsibility is of this form as soon as we remember what, in everyday contexts, we have known all along—namely, that in order to be responsible, an agent must be *sane*. It is not ordinarily in our power to determine whether we are or are not sane. Most of us, it would seem, are lucky, but some of us are not. Moreover, being sane does not necessarily mean that one has any type of power or control an insane person lacks. Some insane people, like JoJo and some actual

political leaders who resemble him, may have complete control of their actions, and even complete control of their acting selves. The desire to be sane is thus not a desire for another form of control; it is rather a desire that one's self be connected to the world in a certain way—we could even say it is a desire that one's self be *controlled by* the world in certain ways and not in others.

This becomes clear if we attend to the criteria for sanity that have historically been dominant in legal questions about responsibility. According to the M'Naughten Rule, a person is sane if (1) he knows what he is doing and (2) he knows that what he is doing is, as the case may be, right or wrong. Insofar as one's desire to be sane involves a desire to know what one is doing—or more generally, a desire to live in the real world—it is a desire to be controlled (to have, in this case, one's *beliefs controlled*) by perceptions and sound reasoning that produce an accurate conception of the world, rather than by blind or distorted forms of response. The same goes for the second constituent of sanity—only, in this case, one's hope is that one's *values* be controlled by processes that afford an accurate conception of the world.<sup>6</sup> Putting these two conditions together, we may understand sanity, then, as the minimally sufficient ability cognitively and normatively to recognize and appreciate the world for what it is.

There are problems with this definition of sanity, at least some of which will become obvious in what follows, that make it ultimately unacceptable either as a gloss on or an improvement of the meaning of the term in many of the contexts in which it is used. The definition offered does seem to bring out the interest sanity has for us in connection with issues of responsibility, however, and some pedagogical as well as stylistic purposes will be served if we use sanity hereafter in this admittedly specialized sense.

#### THE SANE DEEP-SELF VIEW

So far I have argued that the conditions of responsible agency offered by the deep-self view are necessary but not sufficient. Moreover, the gap left open by the deep-self view seems to be one that can be filled only by a metaphysical, and, as it happens, metaphysically impossible addition. I now wish to argue, however, that the condition of sanity, as characterized above, is sufficient to fill the gap. In other words, the deep-self view, supplemented by the condition of sanity, provides a satisfying conception of responsibility. The conception of responsibility I am proposing, then, agrees with the deep-self view in requiring that a responsible agent be able to govern her (or his)

<sup>6</sup> Strictly speaking, perception and sound reasoning may not be enough to ensure the ability to achieve an accurate conception of what one is doing and especially to achieve a reasonable normative assessment of one's situation. Sensitivity and exposure to certain realms of experience may also be necessary for these goals. For the purpose of this essay, I understand "sanity" to include whatever it takes to enable one to develop an adequate conception of one's world. In other contexts, however, this would be an implausibly broad construction of the term.

actions by her desires and to govern her desires by her deep self. In addition, my conception insists that the agent's deep self be sane, and claims that this is *all* that is needed for responsible agency. By contrast to the plain deep-self view, let us call this new proposal the *sane deep-self view*.

It is worth noting, to begin with, that this new proposal deals with the case of JoJo and related cases of deprived childhood victims in ways that better match our pretheoretical intuitions. Unlike the plain deep-self view, the sane deep-self view offers a way of explaining why JoJo is not responsible for his actions without throwing our own responsibility into doubt. For, although like us, JoJo's actions flow from desires that flow from his deep self, unlike us, JoJo's deep self is itself insane. Sanity, remember, involves the ability to know the difference between right and wrong, and a person who, even on reflection, cannot see that having someone tortured because he failed to salute you is wrong plainly lacks the requisite ability.

Less obviously, but quite analogously, this new proposal explains why we give less than full responsibility to persons who, though acting badly, act in ways that are strongly encouraged by their societies—the slaveowners of the 1850s, the Nazis of the 1930s, and many male chauvinists of our fathers' generation, for example. These are people, we imagine, who falsely believe that the ways in which they are acting are morally acceptable, and so, we may assume, their behavior is expressive of or at least in accordance with these agents' deep selves. But their false beliefs in the moral permissibility of their actions and the false values from which these beliefs derived may have been inevitable, given the social circumstances in which they developed. If we think that the agents could not help but be mistaken about their values, we do not blame them for the actions those values inspired.<sup>7</sup>

It would unduly distort ordinary linguistic practice to call the slaveowner, the Nazi, or the male chauvinist even partially or locally insane. Nonetheless, the reason for withholding blame from them is at bottom the same as the reason for withholding it from JoJo. Like JoJo, they are, at the deepest level, unable cognitively and normatively to recognize and appreciate the world for what it is. In our sense of the term, their deepest selves are not fully *sane*.

The sane deep-self view thus offers an account of why victims of deprived childhoods as well as victims of misguided societies may not be responsible for their actions, without implying that we are not responsible for ours. The actions of these others are governed by mistaken conceptions of value that

<sup>7</sup> Admittedly, it is open to question whether these individuals were in fact unable to help having mistaken values, and indeed, whether recognizing the errors of their society would even have required exceptional independence or strength of mind. This is presumably an empirical question, the answer to which is extraordinarily hard to determine. My point here is simply that if we believe they are unable to recognize that their values are mistaken, we do not hold them responsible for the actions that flow from these values, and if we believe their ability to recognize their normative errors is impaired, we hold them less than fully responsible for the relevant actions.

the agents in question cannot help but have. Since, as far as we know, our values are not, like theirs, unavoidably mistaken, the fact that these others are not responsible for their actions need not force us to conclude that we are not responsible for ours.

But it may not yet be clear why sanity, in this special sense, should make such a difference—why, in particular, the question of whether someone's values are unavoidably *mistaken* should have any bearing on their status as responsible agents. The fact that the sane deep-self view implies judgments that match our intuitions about the difference in status between characters like JoJo and ourselves provides little support for it if it cannot also defend these intuitions. So we must consider an objection that comes from the point of view we considered earlier which rejects the intuition that a relevant difference can be found.

Earlier, it seemed that the reason JoJo was not responsible for his actions was that although his actions were governed by his deep self, his deep self was not up to him. But this had nothing to do with his deep self's being mistaken or not mistaken, evil or good, insane or sane. If JoJo's values are unavoidably mistaken, our values, even if not mistaken, appear to be just as unavoidable. When it comes to freedom and responsibility, isn't it the unavoidability, rather than the mistakenness, that matters?

Before answering this question, it is useful to point out a way in which it is ambiguous: The concepts of avoidability and mistakenness are not unequivocally distinct. One may, to be sure, construe the notion of avoidability in a purely metaphysical way. Whether an event or state of affairs is unavoidable under this construal depends, as it were, on the tightness of the causal connections that bear on the event's or state of affairs' coming about. In this sense, our deep selves do seem as unavoidable for us as JoJo's and the others' are for them. For presumably we are just as influenced by our parents, our cultures, and our schooling as they are influenced by theirs. In another sense, however, our characters are not similarly unavoidable.

In particular, in the cases of JoJo and the others, there are certain features of their characters that they cannot avoid *even though these features are seriously mistaken, misguided, or bad*. This is so because, in our special sense of the term, these characters are less than fully sane. Since these characters lack the ability to know right from wrong, they are unable to revise their characters on the basis of right and wrong, and so their deep selves lack the resources and the reasons that might have served as a basis for self-correction. Since the deep selves *we* unavoidably have, however, are sane deep selves—deep selves, that is, that unavoidably *contain* the ability to know right from wrong—we unavoidably do have the resources and reasons on which to base self-correction. What this means is that though in one sense we are no more in control of our deepest selves than JoJo et al., it does not follow in our case, as it does in theirs, that we would be the way we are, even if it is a bad or wrong way to be. However, if this does not follow, it seems to me, our absence of control at the deepest level should not upset us.

Consider what the absence of control at the deepest level amounts to for us: Whereas JoJo is unable to control the fact that, at the deepest level, he is not fully sane, we are not responsible for the fact that, at the deepest level, we are. It is not up to us to *have* minimally sufficient abilities cognitively and normatively to recognize and appreciate the world for what it is. Also, presumably, it is not up to us to have lots of other properties, at least to begin with—a fondness for purple, perhaps, or an antipathy for beets. As the proponents of the plain deep-self view have been at pains to point out, however, we do, if we are lucky, have the ability to revise our selves in terms of the values that are held by or constitutive of our deep selves. If we are lucky enough both to have this ability and to have our deep selves be sane, it follows that although there is much in our characters that we did not choose to have, there is nothing irrational or objectionable in our characters that we are compelled to keep.

Being sane, we are able to understand and evaluate our characters in a reasonable way, to notice what there is reason to hold on to, what there is reason to eliminate, and what, from a rational and reasonable standpoint, we may retain or get rid of as we please. Being able as well to govern our superficial selves by our deep selves, then, we are able to change the things we find there is reason to change. This being so, it seems that although we may not be *metaphysically* responsible for ourselves—for, after all, we did not create ourselves from nothing—we are *morally* responsible for ourselves, for we are able to understand and appreciate right and wrong, and to change our characters and our actions accordingly.

#### **SELF-CREATION, SELF-REVISION, AND SELF-CORRECTION**

At the beginning of this chapter, I claimed that recalling that sanity was a condition of responsibility would dissolve at least some of the appearance that responsibility was metaphysically impossible. To see how this is so, and to get a fuller sense of the sane deep-self view, it may be helpful to put that view into perspective by comparing it to the other views we have discussed along the way.

As Frankfurt, Watson, and Taylor showed us, in order to be free and responsible we need not only to be able to control our actions in accordance with our desires, we need to be able to control our desires in accordance with our deepest selves. We need, in other words, to be able to *revise* ourselves—to get rid of some desires and traits, and perhaps replace them with others on the basis of our deeper desires or values or reflections. However, consideration of the fact that the selves who are doing the revising might themselves be either brute products of external forces or arbitrary outputs of random generation made us wonder whether the capacity for self-revision was enough to assure us of responsibility—and the example of JoJo added force to the suspicion that it was not. Still, if the ability to revise ourselves is not enough, the ability to create ourselves does not seem necessary either. Indeed, when you think of it, it is unclear why anyone should

want self-creation. Why should anyone be disappointed at having to accept the idea that one has to get one's start somewhere? It is an idea that most of us have lived with quite contentedly all along. What we do have reason to want, then, is something more than the ability to revise ourselves, but less than the ability to create ourselves. Implicit in the sane deep-self view is the idea that what is needed is the ability to *correct* (or improve) ourselves.

Recognizing that in order to be responsible for our actions, we have to be responsible for our selves, the sane deep-self view analyzes what is necessary in order to be responsible for our selves as (1) the ability to evaluate ourselves sensibly and accurately, and (2) the ability to transform ourselves insofar as our evaluation tells us to do so. We may understand the exercise of these abilities as a process where by we *take* responsibility for the selves that we are but did not ultimately create. The condition of sanity is intrinsically connected to the first ability; the condition that we be able to control our superficial selves by our deep selves is intrinsically connected to the second.

The difference between the plain deep-self view and the sane deep-self view, then, is the difference between the requirement of the capacity for self-revision and the requirement of the capacity for self-correction. Anyone with the first capacity can *try* to take responsibility for himself or herself. However, only someone with a sane deep self—a deep self that can see and appreciate the world for what it is—can self-evaluate sensibly and accurately. Therefore, although insane selves can try to take responsibility for themselves, only sane selves will properly be accorded responsibility.

#### TWO OBJECTIONS CONSIDERED

At least two problems with the sane deep-self view are so glaring as to have certainly struck many readers. In closing, I shall briefly address them. First, some will be wondering how, in light of my specialized use of the term "sanity," I can be so sure that "we" are any saner than the nonresponsible individuals I have discussed. What justifies my confidence that, unlike the slaveowners, Nazis, and male chauvinists, not to mention JoJo himself, we are able to understand and appreciate the world for what it is? The answer to this is that nothing justifies this except widespread intersubjective agreement and the considerable success we have in getting around in the world and satisfying our needs. These are not sufficient grounds for the smug assumption that we are in a position to see the truth about *all* aspects of ethical and social life. Indeed, it seems more reasonable to expect that time will reveal blind spots in our cognitive and normative outlook, just as it has revealed errors in the outlooks of those who have lived before. But our judgments of responsibility can only be made from here, on the basis of the understandings and values that we can develop by exercising the abilities we do possess as well and as fully as possible.

If some have been worried that my view implicitly expresses an overconfidence in the assumption that we are sane and therefore right about the

world, others will be worried that my view too closely connects sanity with being right about the world, and fear that my view implies that anyone who acts wrongly or has false beliefs about the world is therefore insane and so not responsible for his or her actions. This seems to me to be a more serious worry, which I am sure I cannot answer to everyone's satisfaction.

First, it must be admitted that the sane deep-self view embraces a conception of sanity that is explicitly normative. But this seems to me a strength of that view, rather than a defect. Sanity *is* a normative concept, in its ordinary as well as in its specialized sense, and severely deviant behavior, such as that of a serial murderer or a sadistic dictator, does constitute evidence of a psychological defect in the agent. The suggestion that the most horrendous, stomach-turning crimes could be committed only by an insane person—an inverse of Catch-22, as it were—must be regarded as a serious possibility, despite the practical problems that would accompany general acceptance of that conclusion.

But, it will be objected, there is no justification, in the sane deep-self view, for regarding only horrendous and stomach-turning crimes as evidence of insanity in its specialized sense. If sanity is the ability cognitively and normatively to understand and appreciate the world for what it is, then *any* wrong action or false belief will count as evidence of the absence of that ability. This point may also be granted, but we must be careful about what conclusion to draw. To be sure, when someone acts in a way that is not in accordance with acceptable standards of rationality and reasonableness, it is always appropriate to look for an explanation of why he or she acted that way. The hypothesis that the person was unable to understand and appreciate that an action fell outside acceptable bounds will always be a possible explanation. Bad performance on a math test always suggests the possibility that the testee is stupid. Typically, however, other explanations will be possible, too—for example, that the agent was too lazy to consider whether his or her action was acceptable, or too greedy to care, or, in the case of the math testee, that he or she was too occupied with other interests to attend class or study. Other facts about the agent's history will help us decide among these hypotheses.

This brings out the need to emphasize that sanity, in the specialized sense, is defined as the *ability* cognitively and normatively to understand and appreciate the world for what it is. According to our commonsense understandings, having this ability is one thing and exercising it is another—at least some wrong-acting, responsible agents presumably fall within the gap. The notion of “ability” is notoriously problematic, however, and there is a long history of controversy about whether the truth of determinism would show our ordinary ways of thinking to be simply confused on this matter. At this point, then, metaphysical concerns may voice themselves again—but at least they will have been pushed into a narrower, and perhaps a more manageable, corner.

The sane deep-self view does not, then, solve all the philosophical problems connected to the topics of free will and responsibility. If anything, it

highlights some of the practical and empirical problems, rather than solves them. It may, however, resolve some of the philosophical, and particularly, some of the metaphysical problems, and reveal how intimate are the connections between the remaining philosophical problems and the practical ones.

### DISCUSSION QUESTIONS

1. Explain the deep self views of freedom and responsibility from Frankfurt, Watson, and Taylor, as Wolf describes them.
2. What problems does Wolf raise for the deep self views of Frankfurt, Watson, and Taylor? How could Frankfurt, Watson, and Taylor best respond to Wolf?
3. Explain Wolf's example of JoJo. What point is she trying to make with this example? Does the example succeed in your opinion? Why or why not?
4. Does Wolf's sane deep-self theory really solve the problems raised by her example of JoJo? How or why not?
5. What exactly does Wolf mean by "sanity"? How is her notion related to what people usually mean by sanity and insanity? Are normal people like us sane by Wolf's standards? Why or why not?
6. Wolf distinguishes *actually* understanding and appreciating the world from being *able* to understand and appreciate the world. How can we draw this distinction in practice, such as in a legal trial?
7. Wolf admits, "The sane deep-self view does not, then, resolve all of the philosophical problems connected to the topics of free will and responsibility." Which problems does it *not* resolve?

## A DEFENSE OF FREE WILL SKEPTICISM<sup>2</sup>

by Derk Pereboom

### 1. OUTLINE OF HARD INCOMPATIBILISM.

Spinoza maintained that due to certain general facts about the nature of the universe, we human beings do not have the sort of free will required for moral responsibility. I agree. More exactly, he argues that it is because causal determinism is true that we lack this sort of free will; he is thus a *hard determinist*. By contrast, the position I defend is agnostic about causal determinism. I contend, like Spinoza, that we would not have the sort of free will required for moral responsibility if causal determinism were true, but also that indeterministic theories do not significantly improve the prospects for this sort of free will. Consequently, we need to take seriously the verdict that we lack the sort of free will required for moral responsibility. I call the resulting skeptical view about free will *hard incompatibilism*. In addition, I argue

that a conception of life without this sort of free will need not exclude morality or our sense of meaning in life, and in some respects it could even be beneficial.

## 2. AGAINST COMPATIBILISM.

The case for hard incompatibilism involves arguing against two competing positions. The first of these is *compatibilism*, which claims that free will of the type required for moral responsibility is compatible with determinism. Compatibilists typically maintain, in addition, that we do in fact have this sort of free will. The second is *libertarianism*, which contends that although the sort of free will required for moral responsibility is not compatible with determinism, it turns out that determinism is false, and we do have this kind of free will.

Compatibilists typically attempt to formulate conditions on agency intended to provide an account of what it is to be morally responsible for an action. These conditions are compatibilist in that they allow for an agent to be morally responsible for an action even when she is causally determined to act as she does. For instance, Hume and his followers specify that morally responsible action be caused by desires that flow from the agent's "durable and constant" character, and that the agent not be constrained to act, at least in the sense that the action not result from an irresistible desire. Harry Frankfurt proposes that moral responsibility requires that the agent have endorsed and produced her will to perform the action in the right way. More specifically, she must have a second-order desire—that is, a desire to have a particular desire—to will to perform it, and her will must be her will because she has this second-order desire. John Fischer argues that morally responsible action must result from a rational consideration of the reasons at issue; among other things, the agent must be receptive to the reasons present in a situation, and she must be responsive to them to the degree that in at least some situations in which the reasons are different, she would have done otherwise. Finally, Jay Wallace proposes that moral responsibility requires that the agent have the general ability to grasp, apply, and regulate her behavior by moral reasons. Each of these compatibilists intends for his conditions to be sufficient for an agent's moral responsibility when they are supplemented by some fairly uncontroversial additional necessary conditions, such as the provision that the agent understands that killing is morally wrong.

In my view, the best type of challenge to the compatibilist begins with the intuition that if someone is causally determined to act by other agents, for example, by scientists who manipulate her brain, then she is not morally responsible for that action. This intuition remains strong even if she meets the compatibilist conditions on moral responsibility just canvassed. The following "four-case argument" first of all develops examples of actions that involve such manipulation, in which these compatibilist conditions on moral responsibility are satisfied. Manipulation cases, taken individually, indicate

that it is possible for an agent not to be morally responsible even if the compatibilist conditions are satisfied, and that as a result these conditions are inadequate. But the argument has additional force, by way of setting out three such cases, each progressively more like a fourth scenario which the compatibilist would regard as realistic, in which the action is causally determined in a natural way. An additional challenge for the compatibilist is to point out a relevant difference between any two adjacent cases that would show why the agent might be morally responsible in the later example but not in the earlier one. I argue that this can't be done. So I contend that the agent's non-responsibility generalizes from the first two manipulation examples to the ordinary case.

In each of the four cases, Professor Plum decides to kill Ms. White for the sake of some personal advantage, and succeeds in doing so. We design the cases so that his act of murder conforms to the prominent compatibilist conditions. Plum's action meets the Humean conditions, since for him purely selfish reasons typically weigh heavily—much too heavily as judged from the moral point of view, while in addition the desire that motivates him to act is nevertheless not irresistible for him, and in this sense he is not constrained to act. It fits the condition proposed by Frankfurt: Plum's effective desire (i.e., his will) to murder White conforms appropriately to his second-order desires for which effective desires he will have. That is, he not only wills to murder her, but he also wants to will to do so. The action also satisfies the reasons-responsiveness condition advocated by Fischer: Plum's desires are modified by, and some of them arise from, his rational consideration of the reasons he has, and if he knew that the bad consequences for himself that would result from killing White would be much more severe than they are actually likely to be, he would have refrained from killing her for this reason. In addition, this action meets the condition advanced by Wallace: Plum retains general ability to grasp, apply, and regulate his behavior by moral reasons. When egoistic reasons that count against acting morally are relatively weak, he will usually regulate his behavior by moral reasons instead. This ability provides him with the capacity to revise and develop his moral character over time, as Alfred Mele requires. Now, supposing that causal determinism is true, is it plausible that Plum is morally responsible for his action?

Each of the four cases features different ways in which Plum's murder of White might be causally determined by factors beyond his control.

Case 1: A team of neuroscientists is able to manipulate Professor Plum's mental state at any moment through the use of radio-like technology. In this case, they do so by pressing a button just before he begins to reason about his situation. This causes Plum's reasoning process to be egoistic, which the neuroscientists know will deterministically result in his decision to kill White. Plum does not think and act contrary to character since his reasoning processes are not infrequently egoistic. His effective first-order desire to kill White conforms to his second-order desires. The process of deliberation from which his action results

is reasons-responsive; in particular, this type of process would have resulted in his refraining from killing White in some situations in which the reasons were different. Still, his reasoning is not in general exclusively egoistic, since he often regulates his behavior by moral reasons, especially when the egoistic reasons are relatively weak. He is also not constrained, in the sense that he does not act because of an irresistible desire—the neuroscientists do not induce a desire of this kind.

In Case 1, Plum's action satisfies all the compatibilist conditions we just examined. But intuitively, he is not morally responsible for the murder, because his action is causally determined by what the neuroscientists do, which is beyond his control. Consequently, it would seem that these compatibilist conditions are not sufficient for moral responsibility—even if all taken together.

A compatibilist might resist this conclusion by arguing that although in Case 1 the process resulting in the action satisfies all of the prominent compatibilist conditions, yet Plum's relevant states are directly produced by the manipulators at the time of the action—he is locally manipulated—and this is the aspect of the story that undermines his moral responsibility. In reply, if the neuroscientists did all of their manipulating during one time interval and, after some length of time, the relevant states were produced in him, would he only then be morally responsible? It is my sense that such a time lag, all by itself, would make no difference to whether an agent is responsible.

So let us now consider a scenario in which the manipulation takes place long before the action:

Case 2: Plum is like an ordinary human being, except that neuroscientists have programmed him at the beginning of his life so that his reasoning is frequently but not always egoistic (as in Case 1), with the consequence that in the particular circumstances in which he now finds himself, he is causally determined to engage in the egoistic reasons-responsive process of deliberation and to have the set of first- and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his behavior by moral reasons, but in his circumstances, due to the egoistic character of his reasoning, he is causally determined make his decision. At the same time, he does not act because of an irresistible desire.

Here again, although Plum meets each of the compatibilist conditions, it is intuitive that he is not morally responsible. Thus Case 2 also shows that the prominent compatibilist conditions, either separately or in conjunction, are not sufficient for moral responsibility. Again, it would seem unprincipled to claim that here, by contrast with Case 1, Plum is morally responsible because the length of time between the programming and the action is now great enough. Here also it would seem that he is not morally responsible because he is causally determined to decide and act by forces beyond his control.

Imagine next a scenario more similar yet to an ordinary situation:

Case 3: Plum is an ordinary human being, except that he was causally determined by the rigorous training practices of his household and community in such a way that his reasoning processes are often but not exclusively rationally

egoistic (as in Cases 1 and 2). This training took place when he was too young to have the ability to prevent or alter the practices that determined this aspect of his character. This training, together with his particular current circumstances, causally determines him to engage in the egoistic reasons-responsive process of deliberation and to have the first- and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his behavior by moral reasons, but in his circumstances, due to the egoistic nature of his reasoning processing, he is causally determined make his decision. Here again his action is not due to an irresistible desire.

If a compatibilist wishes to contend that Plum is morally responsible in Case 3, he needs to point to a feature of these circumstances that would explain why he is morally responsible here but not in Case 2. But it seems that there is no such feature. In each of these examples, Plum meets all the prominent compatibilist conditions for morally responsible action, so a divergence in judgment about moral responsibility between these examples will not be supported by a difference in whether these conditions are satisfied. Causal determination by factors beyond his control most plausibly explains the absence of moral responsibility in Case 2, and we are constrained to conclude that Plum is not morally responsible in Case 3 for the same reason.

Thus it appears that Plum's exemption from responsibility in Cases 1 and 2 generalizes to the nearer-to-normal Case 3. Does it generalize all the way to the ordinary case?

Case 4: Physicalist determinism is true—everything in the universe is physical, and everything that happens is causally determined by virtue of the past states of the universe in conjunction with the laws of nature. Plum is an ordinary human being, raised in normal circumstances, and again his reasoning processes are frequently but not exclusively egoistic (as in Cases 1–3). His decision to kill White results from his reasons-responsive process of deliberation, and he has the specified first- and second-order desires. Again, he has the general ability to grasp, apply, and regulate his behavior by moral reasons, and his action is not due to an irresistible desire.

Given that we need to deny moral responsibility in Case 3, could Plum be responsible in this more ordinary case? There would seem to be no differences between Case 3 and Case 4 that could serve to justify the claim that Plum is not responsible in Case 3 but is in Case 4. One distinguishing feature of Case 4 is that the causal determination of Plum's crime is not brought about by other agents. However, the claim that this is a relevant difference is implausible. Imagine a further example that is exactly the same as, say, Case 1 or Case 2, except that Plum's states are induced by a spontaneously generated machine—a machine that has no intelligent designer. Here also Plum would not be morally responsible.

The best explanation for the intuition that Plum is not morally responsible in the first three cases is that his action is produced by a deterministic causal process that traces back to factors beyond his control. Because his action is also causally determined in this way in Case 4, we should conclude that here again he is not morally responsible. So by this argument, Plum's

non-responsibility in Case 1 generalizes to non-responsibility in Case 4. We should conclude that if an action results from any deterministic causal process that traces back to factors beyond the agent's control, then she will lack the control required to be morally responsible for it.

### **3. AGENT-CAUSAL LIBERTARIANISM AND AN OBJECTION FROM OUR BEST PHYSICAL THEORIES.**

A number of different libertarian theories have been proposed, but let us examine the agent-causal version, since it is particularly intuitive and attractive. This view claims that we agent possess a special causal power—a power for an agent, fundamentally as a substance, to cause a decision, and thereby to settle which of a number of competing possible decisions occurs, without being causally determined to do so. But can this position be reconciled with what we would expect given our best physical theories? When an agent makes a free decision, she causes the decision without being causally determined. On the path to action that results from this undetermined decision, changes in the physical world, for example in the agent's brain or some other part of her body, are produced. But if the physical world were generally governed by deterministic laws, it seems that here we would encounter divergences from these laws. For the changes in the physical world that result from the undetermined decision would themselves not be causally determined, and they would thus not be governed by the deterministic laws. For this reason, agent-causal libertarianism is not plausibly reconciled with the physical world's being governed by deterministic laws.

On the standard interpretation of quantum mechanics, however, the physical world is not in fact deterministic, but is rather governed by probabilistic statistical laws. Some philosophers have defended the claim that agent-causal libertarianism can be reconciled with physical laws of this sort. However, wild coincidences would also arise on this suggestion. Consider the class of possible actions each of which has a physical component whose antecedent probability of occurring is approximately 0.32. It would not violate the statistical laws in the sense of being logically incompatible with them if, for a large number of instances, the physical components in this class were not actually realized close to 32% of the time. Rather, the force of the statistical law is that for a large number of instances it is correct to *expect* physical components in this class to be realized close to 32% of the time. Are free choices on the agent-causal libertarian model compatible with what the statistical law leads us to expect about them? If they were, then for a large enough number of instances the possible actions in our class would almost certainly be freely chosen close to 32% of the time. But if the occurrence of these physical components were settled by the choices of agent-causes, then their actually being chosen close to 32% of the time would amount to a wild coincidence. The proposal that agent-caused free choices do not diverge from what the statistical laws predict for the physical components of our actions would run so sharply counter to what we would expect as to make it incredible.

At this point, the libertarian might propose that there actually do exist divergences from the probabilities that we would expect without the presence of agent-causes, and that these divergences are to be found at the interface between the agent-cause and that which it directly affects—an interface which is likely to be found in the brain. The problem for this proposal, however, is that we have no evidence that such divergences occur. This difficulty, all by itself, provides a strong reason to reject this approach.

On the other hand, nothing we've said conclusively rules out the claim that because we are agent-causal, there exist such divergences. We do not have a complete understanding of the human neural system, and it may turn out that some human neural structures are significantly different from anything else in nature we understand, and that they serve to ground agent causation. This approach may be the best one for libertarians to pursue. But at this point we have no evidence that it will turn out to be correct. So given kinds of difficulties we've encountered for libertarianism and compatibilism, it makes sense to take seriously the remaining skeptical option, according to which we do not have the sort of free will required for moral responsibility.

#### 4. HARD INCOMPATIBILISM AND WRONGDOING.

Accepting hard incompatibilism demands giving up our ordinary view of ourselves as blameworthy for immoral actions and praiseworthy for actions that are morally exemplary. At this point one might object that giving up our belief in moral responsibility would have very harmful consequences, perhaps so harmful that thinking and acting as if hard incompatibilism is true is not a feasible option. Thus even if the claim that we are morally responsible turns out to be false, there may yet be weighty practical reasons to believe that we are, or at least to treat people as if they were.

For instance, one might think that if we gave up the belief that people are blameworthy, we could no longer legitimately judge any actions as wrong or even bad, or as right or good. But this seems mistaken. Even if we came to believe that some perpetrator of genocide was not morally responsible because of some degenerative brain disease he had, we would still maintain that his actions were morally wrong, and that it was extremely bad that he acted as he did. So, in general, denying blameworthiness would not at the same time threaten judgments of wrongness or badness, and, likewise, denying praise-worthiness would not undermine assessments of rightness or goodness.

Perhaps treating wrongdoers as blameworthy is often required for effective moral education and improvement. If we resolved never to treat people as blameworthy, one might fear that we would be left with insufficient leverage to reform immoral behavior. Still, this option would have us treat people as blameworthy—by, for example, expressing anger toward them because of what they have done—when they do not deserve it, which would seem *prima facie* morally wrong. If people are not morally responsible for immoral behavior, treating them as if they were would seem to be unfair. However, it is possible to achieve moral reform by methods that would not be threatened

by this sort of unfairness, and in ordinary situations such practices could arguably be as successful as those that presuppose moral responsibility. Instead of treating people as if they deserve blame, the hard incompatibilist can turn to moral admonition and encouragement, which presuppose only that the offender has done wrong. These methods can effectively communicate a sense of right and wrong, and they can issue in salutary reform.

But does this position have resources adequate for contending with criminal behavior? Here it would appear to be at a disadvantage, and if so, practical considerations might yield strong reasons to treat criminals as if they were morally responsible. First of all, if the free will skeptic is right, a retributivist justification for criminal punishment would be unavailable, for it asserts that the criminal deserves pain or deprivation just for committing the crime, while hard incompatibilism denies this claim. And retributivism is one of the most naturally compelling ways to justify criminal punishment.

By contrast, a theory that justifies criminal punishment on the ground that punishment educates criminals morally is not threatened by hard incompatibilism specifically. However, we lack significant empirical evidence that punishing criminals brings about moral education, and without such evidence, it would be wrong to punish them in order to achieve this goal. In general, it is wrong to harm a person for the sake of realizing some good in the absence of impressive evidence that the harm will produce the good. Moreover, even if we had impressive evidence that punishment was effective in morally educating criminals, we should prefer non-punitive ways of achieving this result, if they are available—whether or not criminals are morally responsible.

Deterrence theories have it that punishing criminals is justified for the reason that it deters future crime. The two most-discussed deterrence theories, the utilitarian version and the one that grounds the right to punish on the right to self-defense, are not undermined by hard incompatibilism *per se*. Still, they are questionable on other grounds. The utilitarian theory, which claims that punishment is justified because it maximizes utility (i.e., the quantity of happiness or pleasure minus the quantity of unhappiness or pain), faces well-known challenges. It would seem at times to require punishing the innocent when doing so would maximize utility; in certain situations it would appear to prescribe punishment that is unduly severe; and it would authorize harming people merely as means to the well-being, in this case the safety, of others. The sort of deterrence theory that grounds the right to punish in the right of individuals to defend themselves against immediate threats is also objectionable. For when a criminal is sentenced to punishment he is most often not an immediate threat to anyone, since he is then in the custody of the law, and this fact about his circumstances distinguishes him from those who can legitimately be harmed on the basis of the right of self-defense.

There is, however, a resilient theory of crime prevention that is consistent with hard incompatibilism. This view draws an analogy between the treatment of criminals and the treatment of carriers of dangerous diseases. Ferdinand Schoeman argues that if we have the right to quarantine carriers

of serious communicable diseases to protect people, then for the same reason we also have the right to isolate the criminally dangerous. Notice that quarantining a person can be justified when she is not morally responsible for being dangerous to others. If a child is infected with a deadly contagious virus that was transmitted to her before she was born, quarantine can still be legitimate. Now imagine that a serial killer poses a grave danger to a community. Even if he is not morally responsible for his crimes (say because no one is ever morally responsible), it would be as legitimate to isolate him as it is to quarantine a non-responsible carrier of a serious communicable disease.

Clearly, it would be morally wrong to treat carriers of communicable diseases more severely than is required to protect people from the resulting threat. Similarly, it would be wrong to treat criminals more harshly than is required to protect society against the danger posed by them. Furthermore, just as it would be wrong to quarantine someone whose disease was less than severe, so it would be wrong to lock someone up whose crime was less than severe. In addition, I suspect that a theory modelled on quarantine would not justify measures of the sort whose legitimacy is most in doubt, such as the death penalty or confinement in the worst prisons we have. Moreover, it would demand a degree of concern for the rehabilitation and well-being of the criminal that would alter much of current practice. Just as society must seek to cure the diseased it quarantines, so it would be required to try to rehabilitate the criminals it detains. In addition, if a criminal cannot be rehabilitated, and if protection of society demands his indefinite confinement, there would be no justification for making his life more miserable than needed to guard against the danger he poses.

##### 5. MEANING IN LIFE.

If hard incompatibilism is true, could we legitimately retain a sense of achievement for what makes our lives fulfilled, happy, satisfactory, or worthwhile, and our hold on to our hopes for making these sorts of achievements in our lives? It might be argued that if hard incompatibilism is true, there can be no genuine achievements, for an agent cannot have an achievement for which she is not also praiseworthy. However, achievement, and our hope for achievement, is not as closely connected to praiseworthiness as this objection supposes. If an agent hopes to achieve success in some project, and if she accomplishes what she hoped for, intuitively this outcome would be an achievement of hers even if she is not praiseworthy for it—although at the same time the sense in which it is her achievement may be diminished. For example, if someone hopes that her efforts as a teacher will result in well-educated children, and they do, there remains a clear sense in which she has achieved what she hoped for—even if it turns out she is not praiseworthy for anything she does.

One might think that hard incompatibilism would instill an attitude of resignation to whatever the future holds in store, and would thereby undermine any hope or motivation for achievement. But this isn't clearly right.

Even if what we know about our behavioral dispositions and our environment gives us reason to believe that our futures will turn out in a particular way, it can often be reasonable to hope that they will turn out differently. For this to be so, it may sometimes be important that we lack complete knowledge of our dispositions and environmental conditions. For instance, imagine that someone aspires to become a successful politician, but he is concerned that his fear of public speaking will get in the way. He does not know whether this fear will in fact frustrate his ambition, since it is open for him that he will overcome this problem, perhaps due to a disposition for resolute self-discipline in transcending obstacles of this sort. As a result, he might reasonably hope that he will get over his fear and succeed in his ambition. Given hard incompatibilism, if he in fact does overcome his problem and succeeds in political life, this will not be an achievement of his in as robust a sense as we might naturally suppose, but it will be his achievement in a substantial sense nonetheless.

Still, one might contend that, although the skeptical view leaves room for a limited foundation of the sense of self-worth, this perspective can nevertheless be damaging to our view of ourselves, to our sense of achievement and self-worth. But, in response, first note that our sense of self-worth—our sense that we have value and that our lives are worth living—is to a non-trivial extent due to features not produced by our will, let alone by free will. People place great value on natural beauty, native athletic ability, and intelligence, none of which have their source in our volition. To be sure, we also value efforts that are voluntary in the sense that they are willed by us—in productive work and altruistic behavior, and indeed, in the formation of moral character. However, does it matter very much to us that these voluntary efforts are also *freely* willed?

Consider how someone comes to have a good moral character. It is not implausible that it is formed to a significant degree as a result of upbringing, and moreover, the belief that this is so is widespread. Parents typically regard themselves as having failed in raising their children if they turn out with immoral dispositions, and parents often take great care to bring their children up to prevent such a result. Accordingly, people often come to believe that they have the good moral character they do largely because they were raised with love and skill. But those who come to believe this about themselves seldom experience dismay because of it. People tend not to become dispirited upon coming to understand that their good moral character is not their own doing, and that they do not deserve a great deal of praise or respect for it. By contrast, they often come to feel more fortunate and thankful. Suppose, however, that there are some who would be overcome with dismay. Would it be justified or even desirable for them to foster the illusion that they nevertheless deserve praise and respect for producing their moral character? I suspect that most would eventually be able to accept the truth without incurring much loss. All of this, I think, would also hold for those who come to believe that they do not deserve praise and respect for

producing their moral character because they are not, in general, morally responsible....

#### 7. THE GOOD IN HARD INCOMPATIBILISM.

Hard incompatibilism also promises substantial benefits for human life. Of all the attitudes associated with the assumption that we are morally responsible, anger seems most closely connected with it. Discussions about moral responsibility most often focus not on how we judge morally exemplary agents, but rather on how we regard those who are morally deficient. Examples designed to elicit a strong intuition that an agent is morally responsible most often feature an especially heinous action, and the intuition usually involves sympathetic anger. It may be, then, that our attachment to the assumption that we are morally responsible derives to a significant degree from the role anger plays in our emotional lives. Perhaps we feel that giving up the assumption of responsibility is threatening because the rationality of anger would be undercut as a result.

The kind of anger at issue is the sort that is directed toward a person who is believed to have behaved immorally—it comprises both moral resentment and indignation. Let us call this attitude *moral anger*. Not all anger is moral anger. One type of non-moral anger is directed toward someone because his abilities are lacking in some respect or because he has performed poorly in some situation. We are sometimes angry with machines for malfunctioning. At times our anger has no object. Still, most human anger is moral anger.

Moral anger comprises a significant part of our moral lives as we ordinarily conceive them. It motivates us to resist abuse, discrimination, and oppression. At the same time, expression of moral anger often has harmful effects, failing to contribute to the well-being either of those toward whom it is directed or of those expressing the anger. Often its expression is intended to cause little else than emotional or physical pain. Consequently, it has a tendency to damage relationships, impair the functioning of organizations, and unsettle societies. In extreme cases, it can motivate people to torture and kill.

The realization that expression of moral anger can be damaging gives rise to a strong demand that it be morally justified when it occurs. The demand to morally justify behavior that is harmful is generally a very strong one, and expressions of moral anger are often harmful. This demand is made more urgent by the fact that we are often attached to moral anger, and that we frequently enjoy expressing it. Most commonly we justify expression of moral anger by arguing that wrongdoers deserve it, and we believe that they deserve it because they are morally responsible for what they do. If hard incompatibilism is true, however, justification of this sort is undermined. Yet given the concerns to which expression of moral anger give rise, this may be a good thing.

Accepting hard incompatibilism is not likely to modify our attitudes to the extent that expression of moral anger ceases to be a problem for us. However, moral anger is often sustained and magnified by the belief that

its object is morally responsible for immoral behavior. Destructive moral anger in relationships is nurtured in this way by the assumption that the other is blameworthy. The anger that fuels ethnic conflicts, for example, is almost always fostered by the conviction that a group of people deserves blame for past wrongs. Hard incompatibilism advocates giving up such beliefs because they are false. As a result, moral anger might decrease, and its expressions subside.

Would the benefits that would result if moral anger were modified in this way compensate for the losses that would ensue? Moral anger motivates us to oppose wrongful behavior. Would we lose the motivation to oppose immorality? If for hard incompatibilist reasons the assumption that wrongdoers are blameworthy is withdrawn, the belief that they have in fact behaved immorally would not be threatened. Even if those who commit genocide are not morally responsible, their actions are nonetheless clearly horribly immoral, and a conviction that this is so would remain untouched. This, together with a commitment to oppose wrongdoing, would permit a resolve to resist abuse, discrimination, and oppression. Accepting hard incompatibilism would thus allow us to retain the benefits moral anger can also provide, while at the same time challenging its destructive effects.

## NOTES

<sup>1</sup>From *Responsibility, Character, and the Emotions: New Essays in Moral Responsibility*, ed. Ferdinand Schoeman (Cambridge: Cambridge University Press, 1987).

<sup>2</sup>From *The Truth about the World: Basic Readings in Philosophy*, ed. James Rachels and Stuart Rachels (New York: McGraw-Hill, 2011).

## DISCUSSION QUESTIONS

1. Explain Pereboom's four-case argument. Is it a slippery-slope argument of the kind discussed in Chapter 13? Is it an inference to "the best explanation for the intuition that Plum is not morally responsible in the first three cases" like those discussed in Chapter 9?
2. How is Pereboom's argument related to the challenges of determinism and bypassing that were discussed in the introduction to this chapter?
3. How could Wolf best respond to Pereboom's argument?
4. Can we be justified in punishing people who break criminal laws when they are neither free nor responsible? If so, how? If not, then how should our criminal laws be reformed—if arguments such as Pereboom's show that people are never free or responsible?
5. Can life have meaning for people who are neither free nor responsible?

---

INDEX

---

*Note:* Page numbers followed by letter n indicate footnotes.

- a priori* probability, 216, 244–245  
abduction, 213  
abortion  
    analogical arguments, 392–394  
    begging the question, 325–326  
    conditions of personhood, 388–389  
    conflicting principles, 390–391  
    contraception objection, 417–418  
    fetus, status of, 385–386, 388–390,  
        392–394  
    “future like ours,” 410–411  
    Marquis on, 409–420  
    moral reasoning and, 384–391  
    pro-choice argument, 387–392  
    pro-life argument, 385–387  
    rape and incest, 387–388, 392–393  
    Reagan on, 390, 395  
    self-defense exception, 386–388,  
        392, 490  
    slippery slope arguments, 283,  
        389–391  
    Thomson on, 392–394, 400, 408  
    violinist analogy, 392–394  
*Abortion and the Conscience of the Nation*, 395  
Abraham and Isaac, story of, 462–463  
absurdity, reduction to, 337–340  
abusive assurances, 49  
acceptance of premises, 184–185  
acid rain, 234  
ad hominem arguments  
    assessment of, 310–311  
    definition, 308–309  
    deniers, 309–310  
    dismissers, 310  
    genetic fallacy, 311  
    inconsistency over time, 311  
    self-sealers, 328–332  
    silencers, 309–310  
Adam, 457  
adversarial situation, 68  
advertisements, 77, 320  
affirmative action, 298–299, 364–365  
affirmative propositions, 155  
affirmatives, universal, 155  
affirming the consequent, 428  
affirming a disjunct, 126  
agnostics, 499  
Ahmadinejad, Mahmoud, 79–80  
al Qaeda, 320  
“all,” 157  
“although,” 51  
ambiguity, 83, 291–295. *See also*  
    fallacies of ambiguity  
ampersand, 115  
amphiboly, 293  
analogical reasoning in ethics,  
    392–393  
analogy, arguments from, 204–214  
    abortion and, 392–394  
    cannibalism and, 210–211  
    conclusions, 206  
    explanations, 207  
    God and, 212  
    importance, 205  
    relevance, 205  
    relevant disanalogies, 206  
    in science, 204–205  
analogy with animals argument,  
    414–415  
“and” in propositional logic, 114–115.  
    *See also* conjunctions  
animals argument, analogy with,  
    414–415  
anomalies, 428  
antecedents, 43, 135, 138, 144  
appeal to cases argument, 412–414  
appeals to authority, 10, 64, 314–318

- appeals to emotion, 319–321  
 appeals to fear, 320  
 appeals to pity, 319  
 appeals to popular opinion, 318–319  
 appeals to tradition, 319  
 Aquinas, St. Thomas, 451  
 Archimedes, 9  
 "Argument Clinic," 3  
 argument form. *See also* standard form of arguments  
     definition, 7–8  
     *modus ponens*, 92, 137  
     process of elimination, 92, 125–126  
     in propositional logic, 119–120  
     testing for validity, 128–132, 134  
 argument from interests, 415–416  
 argument from the heap, 278–280  
 argument markers, 41–44, 62–63, 72  
 argument reconstruction, 103  
 "An Argument that Abortion Is Wrong" (Marquis), 409–421  
 argumentative devices, 62–63  
 argumentative performatives, 24, 43, 62  
 arguments, 3, 14  
     combinations of explanations and justifications, 10–13  
     definition, 3–4  
     explanations, 7–12, 196–201, 207  
     formal analysis of, 113–114  
     justifications, 4–7, 196–197  
     patterns in, 44  
     reasons in, 41, 46  
     soundness of, 94, 323  
     standard form of, 45–46, 79–82, 102  
     standards for evaluation, 54–55, 90–95, 180  
     subargument arrangement, 85–90, 102  
     suppressed premises, 96–102  
     validity of, 91–93  
     validity vs. soundness, 94  
 Aristotle, 275  
 arrow, in chain of reasoning, 86–87  
 Articles of Religion, 458  
 assumptions, 8, 11, 227–229, 233  
 assuring, 48–49, 62–63, 69–72, 81–82  
 atheism, 340, 449  
 atom, splitting of, 315  
 audience, 71  
 Austin, J.L., 23–24, 39n1  
 authorities, citing, 48  
 authority, appeals to, 10–11, 307, 314–318  
 availability heuristic, 242–243  
  
*Bakke case*, 364–368, 376–378  
 bargain, 55  
 Bayes, Thomas, 254  
 Bayes's theorem, 253–261  
 Beebe, Andrew, 72–73  
 begging the question, 325–328  
 Behe, Michael J., 426, 441–445  
 benign, remedial purpose, 366  
 Bennett, William, 288  
 best explanation, inferences to, 195–204  
 beyond a reasonable doubt, 354, 381  
 biased sampling, 186–187  
 biconditionals, 146–147  
 Big Bang theory, 451–452  
 Blackmun, Harry, 365  
 Bohr, Niels, 426  
 Bollinger, Lee, 368  
 borderline cases, 276, 279  
 brain activity, as condition of personhood, 388  
 branching tree diagram, 87  
 Brennan, William, 365–368  
 Breyer, Stephen, 370  
*Brothers Karamazov, The* (Dostoevsky), 461  
*Brown v. Board of Education*, 354–355, 364  
 buoyancy, law of, 8  
 burden of proof, 354, 381–382  
 Bureau of Land Management, 74–75, 89–90  
 Burger, Warren, 283  
 Burt, Cyril, 316  
 Bush, George W., 341–342  
 "but," 51  
  
*California v. Carney*, 283  
*Calvin and Hobbes*, 429

- cannibalism, 210–211  
capital punishment  
begging the question, 325  
as cruel and unusual punishment, 105–110  
English parliamentary debate on, 337  
slippery slope arguments, 245  
Carroll, Lewis, 18  
*Catcher in the Rye, The* (Salinger), 307  
cases, appeal to, 412–414  
categorical forms, 154–162  
categorical immediate inferences, 164–165  
categorical logic, 151–174  
categorical propositions, 152–154  
classical theory, 161  
contradictories, 159–160  
existential commitment, 161–162  
four basic forms, 154–159  
immediate inferences, 164–166  
vs. propositional logic, 151–152  
syllogism, 166–175  
translation into basic categorical forms, 156–159  
validity for categorical arguments, 162–175  
categorical propositions, 152–154  
categorical syllogisms, 166–175  
causal explanation, 216  
causal generalizations, 216  
causal prediction, 216  
causal reasoning, 215–238  
background, 227–228  
background assumptions, 11, 99, 191, 227, 229–230, 233  
concomitant variation, 234–238  
conditionals and, 216  
joint test, 222–224  
Legionnaires' disease example, 230–234  
necessary condition test, 217–219, 221–224  
normality, 228–229  
reaching positive conclusions, 226–228  
rigorous testing, 225–226, 229  
sufficient condition test, 217–226  
causal slippery slope arguments, 285–289  
certainty, 244  
chain of reasoning, 86–87  
chances. *See* probability  
changing the subject, 32  
*Chicago Tribune*, 186  
choices. *See* decision procedure;  
decisions under ignorance;  
decisions under risk; decisions  
under uncertainty  
Churchill, Winston, 41  
circular reasoning, 324, 327  
circularity, 323–324  
civil law, 355  
claims, universal, 334  
clarification of terms in argument reconstruction, 83–85, 102  
classical theory  
contradictories in, 166  
existential commitment, 161–162  
immediate inferences, 164  
vs. modern logic, 174  
clerk hire allowance, 59–79, 285  
Clinton, Bill, 74–75, 333  
close analysis, 59–102  
argument markers, 62–63, 72, 86  
argument reconstruction, 102  
argumentative performatives, 62  
assuring, 81  
audience, 71  
discounting, 62–63  
evaluative language, 69, 72  
guarding, 62, 64–65, 67, 81–82  
labeling argumentative devices, 62  
purpose, 71  
rhetorical devices, 62  
standpoint, 71  
cognitive psychology, 241–242  
Colbert, Stephen, 298, 305, 311, 320–321, 342, 349n1, 426, 448n2  
Columbia, 79–80  
combinations, 250  
commitment, level of, 49  
common knowledge, 229–230  
common law, 357  
composition, fallacy of, 344  
conceptual slippery slope arguments, 280–283  
conclusion markers, 42–43

- conclusions  
 argument markers, 42–43  
 argument reconstruction, 102  
 arguments from analogy, 204–213  
 arranging subarguments, 85–90  
 clarification of terms, 83  
 in deep analysis, 79–81  
 definition, 3  
 explanations and, 7, 10–11  
 fallacies of relevance, 307, 310–311  
 fallacies of vacuity, 323–326  
 fallacies of vagueness, 275, 279–282  
 false, 99, 119–120, 128, 130  
 guarding, 49–50  
 in inductive vs. deductive  
     arguments, 179–183, 185  
 inferences to the best explanation, 196–197, 199  
 justifications and, 4–6, 10–11  
 positive, 226–228  
 probability and, 239–240, 244  
 refutation, 333–334, 337–338, 343, 346  
 soundness and, 94  
 in standard form, 46  
 statistical applications, 189–191  
 suppressed premises and, 96–98  
 syllogisms and, 166–174  
 truth and, 93  
 universal, 183  
 validity and, 91–94, 144, 162–164
- concomitant variation, 234–238  
 conditional probability, 247, 257  
 conditionals, 43–45, 134–150  
     biconditionals, 146–147  
     causal reasoning and, 216–217  
     indicative, 44, 64, 135, 145  
     logical language and everyday  
         language, 142–148  
     probability and, 247, 257  
     propositional, 43–45, 135, 137,  
         140, 145  
     subjunctive, 135  
     truth tables for, 136–142  
 conflicts of moral principles, 384,  
     390–392  
 conjunctions, 52, 114–129, 142,  
     247–248  
 connectives. *See* propositional  
     connectives
- Conoco, Inc., 74–76, 89–90  
 consequents, 43, 135  
 conservative explanations,  
     199–200  
 considered judgment argument,  
     411–412  
 Constitution, 105–108, 356–357,  
     361–362, 426  
*Consumer Reports*, 204  
 context, crucial, 200–201  
 contextual definitions, 305n4  
 contingent facts, 97–99  
 continuum, 280  
 contraception objection, 417–418  
 contracts, law of, 28  
 contradictions, 159–160  
 convention, and language,  
     17–19  
 conversational acts, 28–35  
     and evaluative language, 54  
     as a practical activity, 28  
     rules of, 31–33. *See also*  
         conversational rules  
 conversational exchange, 28  
 conversational implications, 33–35,  
     36, 44, 52  
 conversational rules, 39n6. *See also*  
     manner, rule of; quality, rule  
     of; quantity, rule of; relevance,  
     rule of  
     Grice on, 33, 36–37, 307  
     rhetorical devices and, 36–37  
     specific rules, 31–33  
     violations of, 80, 122  
 conversion, 164–166  
 Cooperative Principle, 31  
 correlations, 234–236  
 counterevidence, 201  
 counterexamples, 336–337  
 Craig, William Lane, 450–456, 458  
 creationism, 332. *See also* evolution,  
     theory of  
 crimes, worst of, argument,  
     411–412  
 criminal law, 354  
 criminal trials, 10–13  
 cruel and unusual punishment,  
     105–108  
 Curley, Edwin, 450, 456–463

- Daalen, D.H. van, 454  
Darwin, Charles, 332, 427–446. *See also* evolution, theory of; natural selection, theory of  
*Darwin's Black Box* (Behe), 442, 444  
Davies, Paul, 453  
Davies, P.C.W., 452  
death penalty. *See* capital punishment  
decision procedure, 140  
decisions under ignorance, 268–272  
decisions under risk, 268  
decisions under uncertainty, 268  
deductive arguments  
categorical logic, 151–175  
definition, 93–94  
vs. inductive arguments, 179–183  
legal reasoning and, 355  
propositional logic, 111–150  
refutation by parallel reasoning, 343–349  
sufficient and necessary conditions, 220–221, 227–228  
deep analysis, 79–110  
arrangement of subarguments, 85–90  
of capital punishment, 105–108  
chain of reasoning, 86–88  
clarification of terms, 83  
dissection of argument, 83–85  
guarding, 81–82  
linguistic principles, 99–100  
reconstruction method, 102–104  
standard form in, 79–82  
suppressed premises, 96–104  
deep counterexample, 335  
defeasible, 181, 221, 198  
“Defense of Abortion, A” (Thomson), 392–393  
Defense of Marriage Act, 80  
definitions, 299–305  
contextual, 305n4  
dictionary, 300  
disambiguating, 293, 300–301  
precising, 301–302  
stipulative, 301  
systematic or theoretical, 302–303  
degenerate cases, 175n1, 175n2  
deniers, 309–310  
denying the antecedent, 138–139  
depth of explanation, 198  
descriptive terms, 56  
Dewey, Thomas, 186–187  
diagnostic test, 258–259  
dictionary definitions, 300  
diminishing marginal utility, 267  
diminishing marginal value, 267  
Dirksen, Everett, 281  
disagreements, moral, 383–384  
disambiguating definitions, 293, 300–301  
disappearing guard, 51  
disaster avoidance rule, 271  
discounting, 48, 51–53, 62–63, 65–66, 69–70  
discounting, universal, 330  
discounting connectives, 51–52  
discrimination, law of, 361–380  
application of the Constitution, 362–363  
*Bakke* case, 364–369, 375–376  
*Brown v. Board of Education*, 354–355, 364  
Equal Protection Clause, 361–362  
*Gratz v. Bollinger*, 375–380  
*Grutter v. Bollinger*, 369–374, 378  
*Korematsu v. United States*, 363  
*Plessy v. Ferguson*, 340, 357–363  
*Richmond v. J.A. Croson Co.*, 368  
*Strauder v. West Virginia*, 362  
strict scrutiny test, 363–364, 381–382  
*United Steelworkers of America v. Weber*, 368  
*U.S. v. Cruikshank*, 362  
*Yick Wo v. Hopkins*, 357–362  
discriminatory effect, 359–360  
discriminatory motive, 360  
disjunction, 122, 126, 142  
disjunction with exclusivity, 132–134, 248–250  
disjunctive syllogism, 125  
dismissers, 310  
dissection of the argument, in deep analysis, 83–85  
distress sale policy, 368  
dominance, rule of, 269  
domino arguments, 285  
Dostoevsky, Fyodor, 461

- Downey, Michael, 210–212  
 Doyle, Arthur Conan, 213
- Eddington, A.S., 439  
 Edwards, John, 80  
 Ehrman, Bart, 458n  
 Einstein, Albert, 315–316, 426, 428  
 either-or sentences, 44, 122. *See also*  
     disjunction  
 elimination, process of, 92, 125–126  
 emotion, appeals to, 319–321  
*Enquiry Concerning Human  
     Understanding* (Hume), 193  
 ensoulment, as condition of  
     personhood, 388  
 enthymeme, 97  
 Equal Exchange Coffee, 76  
 equal probabilities, 269–270  
 Equal Protection Clause,  
     361–362  
 equality, 383–384, 416–417  
 equivocation, fallacy of,  
     295–299, 301  
 Establishment Clause, 426  
 evaluative language, 54–57  
     in close analysis, 62, 68–70  
     vs. descriptive terms, 56  
     standards in, 180  
     terms, 55–56  
     used to perform conversational  
         acts, 54  
     used to perform speech acts, 54  
 “even if” in arguments, 52  
 “even though” in arguments, 52  
 everyday language, 142–148  
 evidence, 354  
 evolution, theory of, 338, 425,  
     427–440, 461, 461n. *See also*  
     natural selection, theory of  
 exaggeration, 36  
 excess verbiage, trick of, 80  
 exclusive disjunction, 132–134,  
     248–250  
 excuses, 12, 69  
 existential commitment, 161–162  
 expected monetary value,  
     263–266  
 expected overall value, 266–268
- expected utility, 267  
 explanations, 7–10  
     in arguments from analogy,  
         207–208  
     causal, 216  
     conservative, 199–200  
     definition, 7  
     depth of, 198  
     example of, 10–13  
     excuses as, 12, 69  
     inferences to the best explanation,  
         196–201  
     vs. justification, 7  
     modesty of, 199  
     power of, 198–199  
     scientific reasoning and, 423  
     simplicity of, 199  
 explicit performatives, 23, 26  
 expressing emotion, 54  
 extent, 49
- factors, weighing, 394–395  
 fairness, 391  
 fairness slippery slope arguments,  
     282–284  
 fallacies of ambiguity, 291–305  
     ambiguity, 83, 291–295  
     definitions, 299–305  
     disambiguating, 293, 300–301  
     equivocation, 295–299, 301  
     semantic ambiguity, 293  
     syntactic ambiguity, 293  
 fallacies of probability, 239–243. *See  
     also* probability  
 gambler’s fallacy, 239–241  
 heuristics, 241–243  
 fallacies of relevance, 275–289  
     ad hominem arguments, 308–314,  
         328–332  
     appeals to authority, 10–11, 64,  
         314–318  
     appeals to emotion, 319–321  
     appeals to fear, 320  
     appeals to pity, 319–320  
     appeals to popular opinion,  
         318–319  
     appeals to tradition, 319  
     red herring, 80, 308

- fallacies of vacuity, 323–332  
begging the question, 324–328  
circularity, 323–324  
self-sealers, 328–332
- fallacies of vagueness, 275–289  
heaps, 278–280  
slippery slopes. *See* slippery slope arguments  
uses of unclarity, 275
- fallacy of affirming the consequent, 424
- fallacy of composition, 344
- fallacy of hasty generalization, 185
- false alarms, 254, 258
- false dichotomies, 341–342
- false negatives, 258
- false positives, 254, 258
- falsifiable, 199
- falsification of data, 316–317
- fear, appeals to, 320
- Federal Communications Commission, 368
- Feinberg, J., 420
- Feldman, F., 417
- fetus, status of, 385–386, 388–390, 392–394. *See also* abortion
- First Amendment, 426
- Five Minute Whodunits* (Smith), 108–110
- “Five Reasons to Believe in God” (Craig), 450–456
- Flew, Antony, 331–332
- FLO theory, 410–418  
analogy with animals argument, 414–415  
appeal to cases argument, 412–414  
argument from interests, 415–416  
considered judgment argument, 411–412  
contraception objection, 417–418  
future like ours, 410–411  
potentiality objection, 415  
problem of equality, 416–417  
worst of crimes argument, 412
- following from, 91, 113. *See also* validity
- “for” in arguments, 42
- formal analysis of arguments, 113–114
- forms, argument. *See* argument form
- forms, categorical, 154–162
- Fourteenth Amendment, 361–362
- Franklin, Benjamin, 96–97
- Fraser, David W., 238n2
- fraternities, 72–74
- fraud, 317
- free will, 460–461
- freedom, 460–461
- Freud, Sigmund, 312, 329–330
- fundamental rights, 363
- Furman v. Georgia*, 106
- “future like ours”. *See* FLO theory
- Gallup poll, 186–187
- gambler’s fallacy, 239–241
- “garden path sentences,” 20
- Geller, Uri, 315
- general conditional, 216
- generalizations, 179–193  
causal, 216  
induction vs. deduction, 179–183  
statistical, 183–188  
statistical applications, 189–193
- genetic code, as condition of personhood, 388–389
- genetic fallacy, 311–314
- Gillie, Oliver, 317
- Gilovich, Thomas, 261n3
- Ginsburg, Ruth Bader, 369–370, 378
- “Gnawing Question Is Answered, A” (Downey), 210–212
- God. *See also* religious reasoning  
Abraham and Isaac story, 462  
agnostics on, 449  
arguments from analogy, 212–213  
atheists on, 340  
complex order of universe, 452  
existence of, 450–455  
Flew on, 331–332  
Job and, 314  
as known and experienced, 455  
objections to belief in, 456–460  
objective moral values, 453  
origin of the universe, 451–452  
predestination, 457, 459, 462  
resurrection of Jesus, 454
- going upstairs, 330
- Gordon, Mary, 395–396
- Gould, Stephen Jay, 317

- grace, doctrine of, 459  
 gradualism, 389, 420n2  
 grammatical conventions, 21  
 grammatical rules, 20  
 Grand Staircase-Escalante National Monument, 74, 89–90  
*Gratz v. Bollinger*, 375–380  
 “Greeks Should Be Co-Ed” (Beebe), 72–74  
 Grice, Paul  
     on conversational rules, 33, 36–37, 307  
     on Cooperative Principle, 31–32  
     “Logic and Conversation,” 39n6  
     on validity, 144  
*Grutter v. Bollinger*, 369–375, 378  
 guard, disappearing, 51  
 guarding, 48–51  
     in close analysis, 62, 65–68, 70  
     in deep analysis, 81–82  
 gun rights, 344–346
- Haeckel, Ernst, 429–430  
 Harlan, John, 360–361  
 Harman, Gilbert, 213n1  
 hasty generalization, fallacy of, 185  
 Hawking, Stephen, 452  
*Hawkins v. McGee*, 28  
 heaps, 278–280  
 hedging, 51  
 hell, 457–459  
 heuristics, 241–243  
 Hick, John, 455  
 hidden premises, 100  
 Hilbert, David, 451  
 Hill, Thomas E. Jr., 298, 305  
 Holmes, Sherlock, 197–198  
 homosexuality, 299  
 horrors, parade of, 67, 285  
 horseshoe, 137  
 House of Representatives, 59–72  
 “however” in arguments, 52  
 Hoyle, Fred, 451–453  
 Hume, David, 193  
 Hussein, Saddam, 13–15  
 Huxley, Henry, 429–430  
 Huxley, Thomas, 338  
 hyperbole, 36
- hypothesis, 198  
 hypothetical cases, 360  
 hypothetical syllogism, 140
- identical twins, 316  
 ideologies, as self-sealing, 329  
 “if and only if” sentences, 147  
 “if not” in arguments, 148  
 if-then statements, 43–45, 135. *See also* conditionals  
 ignorance, decisions under, 268–272  
 illocutionary acts, 39n4  
 immediate inferences, 164–166  
 impersonal normative justification, 5  
 importance, in arguments from analogy, 205  
 imprecision, 276  
 incest, 387–388. *See also* abortion  
 inconsistency over time, 311  
 indefeasible, 181  
 independence, 250–251  
 indicative conditionals, 44, 64, 135, 145  
 inductive arguments, 93, 180, 221  
 inductive generalizations, 179–183  
 inductive inference, 180, 191  
 inductive reasoning, 193  
 inductive standards  
     arguments from analogy, 204–213  
     causal reasoning, 215–238  
     choices, 263–272  
     vs. deductive arguments, 179–183  
     defeasibility, 191  
     induction vs. deduction, 179–183  
     inference to the best explanation, 195–204  
     legal reasoning and, 360  
     Legionnaires’ disease example, 232–234  
     opinion polls, 183–184  
     probability, 239–261  
     sampling, 185–188  
     statistical applications, 189–193  
     statistical generalizations, 183–188  
     strength as, 180–182  
     sufficient and necessary conditions, 220–222, 227–228

- inferences  
arguments from analogy, 204–213, 358  
to the best explanation, 195–204
- initial conditions, 8
- innocent until proven guilty, 381
- insinuations, 50
- insufficient reason, rule of, 269
- intelligence, inheritance of, 316–317
- intelligent design, 426, 438. *See also* creationism
- interests, argument from, 415–416
- Iran, 79–80, 90
- Iraq, 13–15
- Iraq war, 341
- irony, 36, 69
- Isaac and Abraham, story of, 462
- Japanese Americans, 363
- Jastrow, Robert, 453
- Jesus Christ, 454
- Job, 314
- Johnson, Luke, 455
- joint test, 222–224
- Jones, John E. III, 426
- Journal of Molecular Evolution*, 437
- justifications, 4–7, 10–13, 196–197
- Kahneman, Daniel, 261n1
- Kamin, Leon, 317
- Kennedy, Anthony, 373, 375
- Kenny, Anthony, 451
- Kitcher, Philip, 426, 440–448
- Kitzmiller v. Dover Area School District*, 426
- knowledge, 98, 228–229, 231–232, 300, 425, 444
- Korematsu v. United States*, 363
- Kremer, Jacob, 454
- Kuhn, Thomas, 448n1
- Kyl, Representative, 59–72, 87–88, 285
- Landon, Alf, 186–187
- Land's End advertisement, 77
- language, 17–39, 41–57. *See also* arguments; evaluative language  
argument markers, 41–44, 62–63, 70  
assuring, 48–49, 62–66, 69–70, 81–82  
convention and, 17–19  
conversational acts, 28–35  
discounting, 48, 51–53, 62–63, 65–66, 69–70  
evaluative, 54–57  
guarding, 48–51, 62, 65–68, 70, 81–82  
if-then statements, 43–45, 135  
linguistic acts, 17, 19–22, 26, 39n1  
logical vs. everyday, 142–148  
of probability, 243–244  
rhetorical devices, 36–38, 62, 353  
speech acts, 22–28
- large numbers, law of, 240
- legal reasoning, 353–382  
burden of proof, 354, 381–382  
components of, 354–361  
law of discrimination, 361–380  
questions of fact, 354–355, 383  
questions of law, 355–361
- Legionnaires' disease example, 230–234
- legislative intentions, 356
- legitimate and compelling state interest, 363
- level of commitment, 49–50
- Lincoln, Abraham, 19
- linguistic acts, 17, 19–22, 26, 39n1
- linguistic principles, 99–100
- Literary Digest*, 186–187
- "Living with Darwin" (Kitcher), 440–447
- locutionary act, 39n1
- logic. *See* categorical logic; propositional logic
- "Logic and Conversation" (Grice), 39n6
- logical language, 142–145
- logically equivalent, 165
- London, 320
- lotteries, 263–264, 266–267
- Ludemann, Gerd, 455
- lung cancer, 237

- Mackie, J.L., 453  
 major premise, 167  
 major term, 167  
 manner, rule of  
     definition, 32, 34  
     violations of, 80  
 marginal utility, 267  
 marginal value, 267  
 Marquis, Don, 409–420  
 Marshall, Thurgood, 365  
 Marxism, 312, 330  
 material conditional, 137  
 Matthews, Chris, 90  
 maximax rule, 270  
 maximin rule, 270–271  
 McCain, John, 90  
 McDade, Joseph E., 238n2  
 "Mere Matter of Deduction, A"  
     (Smith), 108–109  
 metaphors, 36–37  
 Method of Concomitant  
     Variation, 235  
*Metro Broadcasting v. FCC*, 368  
 Mill, John Stuart, 216–217, 235,  
     296, 328  
 Miller, Zell, 320  
 Mill's methods, 216–217  
 minor premise, 167  
 minor term, 167  
*Misquoting Jesus* (Ehrman), 458n  
 misses, 258  
 modal logic, 144  
 modern logic, 161–162  
 modesty of explanation, 199  
*modus ponens*, 92  
 "Molecular Machines: Experimental  
     Support for the Design Infer-  
     ence" (Behe), 427–440  
 monetary value, expected, 263–266  
 Monty Python, 3  
 moral beliefs, 356, 453  
 "Moral Choice, A" (Gordon), 395–396  
 moral disagreements, 383–384  
 moral principles, 383–384  
 moral reasoning, 383–421  
     abortion. *See* abortion  
     analogical reasoning in ethics,  
         392–394  
     moral disagreements, 383–384  
     problem clarification, 384–385  
     weighing factors, 394–395  
 Morgenbesser retort, 335  
 "must" in arguments, 64  
 mutually exclusive, 248  
 narrative, 7  
 National Rifle Association (NRA),  
     344–346  
 natural selection, theory of, 332,  
     426–440. *See also* evolution,  
     theory of  
 Neanderthals, 210–211  
 necessary condition test (NCT),  
     216–217, 224–228  
     application of, 231–234  
     causal reasoning, 217–219, 221–222  
 necessary conditions, 216–219  
 negation, 122–125, 246–247  
 negative propositions, 155  
 negatives, universal, 155  
 "nevertheless" in arguments, 52  
*New Merriam-Webster Pocket  
     Dictionary*, 291  
 Newton, Isaac, 303, 427–428  
 Nietzsche, Friedrich, 453  
 nonpropositional conjunction, 117  
 normality, 228–230  
 North Korea, 318–319  
 "not always" in arguments, 175n1  
 "not" in arguments, 123–124, 175n1  
 nuclear power, 315–316  
 nuclear weapons, 13–15, 79–80, 139  
 obfuscation, 275  
 objections, raising, 68  
 objective moral values, 453  
 Occam's razor, 199  
 O'Connor, Sandra Day, 283, 368,  
     370, 375  
 oil drilling, 74–76  
 "only if" in arguments, 175n2  
 opinion polls, 183–184  
 "or" in arguments, 122, 132  
*Origin of Species, The* (Darwin),  
     427–429, 432. *See also* natural  
     selection, theory of

original sin, 457  
overall value, expected, 266–268  
overstatement, 36  
*Oxford American Dictionary*, 300

Paley, William, 212–213, 214n8  
parade of horrors, 67, 285  
parallel reasoning, refutation by,  
    343–349  
particular affirmatives, 155  
particular negatives, 155  
particular propositions, 155  
patterns in arguments, 44  
Paul, St., 457n, 458n  
Peirce, Charles Sanders, 213n1  
performative verbs, 26–27  
performatives, 23–27  
permutations, 250–251  
personal justification, 5  
personhood, conditions of, 388–389  
*petitio principii*, 328  
philosophical reasoning, 465–494  
    computers' ability to think, 465–493  
    God and, 471  
    “Defense of Free Will Skepticism,  
        A” (Pereboom), 483–494  
    “Sanity and The Metaphysics  
        of Responsibility,” (Wolf),  
        469–483  
“Piece of ‘God’s Handiwork’, A”  
    (Redford), 74–76, 89–90  
pity, appeals to, 319–320  
*Plessy v. Ferguson*, 340, 357–362  
plus sign, 88  
polls, opinion, 183–184  
popular opinion, appeals to, 318–319  
positive conclusions, 226–228  
positive correlation, 235–236  
potentiality objection, 415  
Potter, Stewart, 106  
Powell, Colin, 13  
Powell, Lewis, 365–368, 376, 379  
power of explanation, 198–199  
precedents, 357–361  
precising definitions, 301–302  
predestination, 457, 459–462  
predicate nominative, 164  
predicate term, 164, 167

prediction, 8, 216, 196, 424  
preferential treatment quota, 368  
prejudice, 187–188  
premises  
    acceptance of, 184–185  
    arrangement of subarguments,  
        85–89  
    begging the question, 324–327  
    circularity and, 323–324  
    clarification of terms, 83–85  
    in close analysis, 64  
    in deep analysis, 79–82  
    definition, 3–4  
    equivocation and, 295–297  
    false conclusions and, 91–93  
    guarding and, 49–51  
    hidden, 100  
    major, 167  
    minor, 167  
    reconstruction of arguments,  
        102–104  
    refutation, 333–334, 343–346  
    rejection of, 6–7  
    soundness and, 94  
    standard form and, 7–8, 46  
    suppressed, 96–104  
    validity and, 91–94, 114, 119–120,  
        128–131, 142–144  
    in Venn diagrams, 162–164  
preponderance of evidence,  
    354, 381  
prescribing action, 54  
*prima facie* case, 381  
probabilistic version of statistical  
    syllogism, 193n1  
probability, 239–261, 263–272  
    *a priori*, 216, 244–246  
    Bayes’s theorem, 253–261  
    conditionals, 247, 257  
    conjunction, 52, 114–121, 142,  
        247–248  
    in choices, 263–272  
    disjunction with exclusivity,  
        132–134, 248–249  
    equal probabilities, 269–270  
    gambler’s fallacy, 239–241  
    heuristics, 241–243  
    language of, 243–244  
    laws of, 246–253

- negation, 122–125, 246–247  
 series with independence, 250–251
- probability phrases, 49
- probability rules, 246–253  
 permutations and combinations, 250–251  
 of conjunctions, 247–248  
 of disjunctions, 248–249  
 of negations, 246–247  
 in a series, 249–250
- probability scale, 244
- problem clarification, 384–386
- problem of equality, 416–417
- process of elimination, 92, 125–126
- pro-choice argument, 387–392. *See also* abortion
- pro-life argument, 385–387. *See also* abortion
- propositional conditionals, 43–45, 135, 137, 140, 145
- propositional conjunction, 117
- propositional connectives, 114–134  
 conjunction, 52, 114–121, 142, 247–248  
 definition, 114  
 disjunction, 122, 126, 142  
 exclusive disjunction, 132–134, 248–249  
 negation, 122–125, 246–247  
 truth-functional connectives, 126–128  
 truth-functionally equivalent, 133–134  
 validity testing, 128–132, 144, 163
- propositional form, 115–116
- propositional logic, 113–150  
 argument form in, 119–120  
 vs. categorical logic, 151–152  
 conditionals, 43–45, 134–150  
 formal analysis of arguments, 113–114  
 validity and, 119–121
- propositional variables, 151
- propositions, 155
- psychic powers, 315
- psychoanalytic theory, 329
- purpose, 71
- quality, rule of  
 contingent facts, 97–99  
 definition, 31  
 and rhetorical devices, 36, 37
- quantity, rule of  
 contingent facts, 97  
 definition, 31  
 linguistic principles, 99–100  
 and rhetorical devices, 36–37  
 violations of, 80, 122, 144–145, 147
- questions of fact, 354–355, 383
- questions of law, 355–361
- quotas, 368
- racial segregation, 340
- raising objections, 68
- Randi, James, 315
- range of moral principles, 384
- rape, 387–388, 392–393. *See also* abortion; violinist analogy, and abortion
- ratio decidendi*, 358
- rationality, as condition of personhood, 388–389
- rational-relation test, 362–363, 381
- Reagan, Ronald, 139, 390, 395
- reason markers, 42–43
- reasoning, chain of, 86–88
- reasons. *See* premises
- reconstruction method, 102–103
- red herring, 80, 308
- Redford, Robert, 74–76, 89–90
- reductio ad absurdum, 337–340
- reference class, 189, 190–192
- refutation, 333–349  
 counterexamples, 335–337  
 definition, 333  
 false dichotomies, 341–342  
 forms of, 333  
 by parallel reasoning, 343–349  
 reductio ad absurdum, 337–340  
 straw man, 52, 64–65, 341–342
- Regents of the University of California v. Bakke*, 364–369, 375–376
- Rehnquist, William, 365, 373
- Reid, Thomas, 214n6
- relevance, fallacies of. *See* fallacies of relevance

- relevance, in arguments from analogy, 205
- relevance, rule of
- contingent facts, 97
  - and conversational implications, 34–35
- definition, 32
- fallacies and, 307
- and rhetorical devices, 36–37
- violations of, 80, 142, 144
- relevant disanalogies, 206
- religious reasoning, 449–463.
- See also* God
- Abraham and Isaac story, 462
  - agnosticism, 449
  - complex order of universe, 452
  - existence of God, 450–456
  - God as known and
    - experienced, 455
  - grace, 458–459
  - hell, 457–458
  - objections to belief in God, 455–460
  - objective moral values, 453
  - origin of the universe, 451–452
  - original sin, 457
  - predestination, 457, 459–462
  - resurrection of Jesus, 454
- repetition, 80–81
- representative heuristic, 241
- resurrection of Jesus, 454
- reverse discrimination, 364. *See also* affirmative action
- rhetorical devices
- in close analysis, 62–63, 68
  - conversational rules, 36–38
- Richmond v. J.A. Croson Co.*, 368
- rigorous testing, 225–226, 234
- risk, decisions under, 268
- Romney, Mitt, 79–80
- Roosevelt, Franklin, 186–187
- rules, of conversational acts, 31–33
- Ruse, Michael, 453
- Rutherford, Ernest Lord, 315
- Salinger, J.D., 321n1
- sample size, 185–186
- sampling, biased, 186–187
- “Sanity and The Metaphysics of Responsibility,” (Wolf), 469–483
- sarcasm, 36
- Satan, 314
- Scalia, Antonin, 373, 375
- Schwarzenegger, Arnold, 97–99
- science, falsification of data in, 316–317
- scientific reasoning, 423–448. *See also* creationism, evolution, theory of; intelligent design; natural selection, theory of
- explanations and, 423
  - scientific revolutions, 425–427
  - standard science, 423–425
- scientific revolutions, 425–427
- scope of moral principles, 384
- screening test, 258
- search warrant, 283
- Sedgwick, John B., 321n4
- “seem” in arguments, 65
- segregation, 340, 363–364
- self-defense exception, 386–388.
- See also* abortion
- selfishness, 330
- self-sealers, 328–332
- semantic ambiguity, 293
- semantic conventions, 21
- semantic rules, 20
- sentience, as condition of personhood, 388
- separate-but-equal doctrine, 363–364
- September 11 terrorist attacks, 320
- series with independence, 250–251
- “Seven Deadly Objections to Belief in the Christian God” (Curley), 456–460
- shallow counterexample, 335
- silencers, 309–310
- “Silver Blaze” (Doyle), 197–198
- similes, 36
- simplicity of explanation, 199
- sir, 55
- “since” in arguments, 42, 86
- slippery slope arguments, 280–289, 302
- abortion and, 283, 389–392
  - causal slippery slope arguments, 285–289

- conceptual slippery slope arguments, 280–282
- domino arguments, 285
- fairness, 282–284
- parade of horrors, 67, 285
- Smith, Stan, 108–110
- smoking, 237
- “so” in arguments, 44, 86
- socialism, 85
- sophistry, 67
- sorites argument, 278
- sororities, 72–74
- soundness, 94
- specialized knowledge, 229
- speech acts, 22–28
- evaluative language and, 54–57
  - fail to come off, 27
  - vs. linguistic acts, 26
  - performatives, 23–27
  - types of, 26–28
  - void, 27
- Spinoza, Benedict de, 450
- Squad Helps Dog Bite Victim and Other Flubs from the Nation’s Press*, 294–295
- standard form of arguments, 7–8, 45–46, 79–82. *See also* argument form
- standard science, 423–425
- standpoint, 59, 71
- Stanwick, Thomas P., 108–109
- stare decisis*, 357
- state action, 362
- state interest, 363
- statistical applications, 189–193
- statistical generalizations, 183–188
- acceptance of premises, 184–185
  - biased sampling, 186–187
  - prejudice/stereotypes, 187–188
  - sample size, 185–186
- statistical syllogism, 193n1. *See also* statistical applications
- statutes, 355–356
- Stewart, John, 288
- Stewart, Potter, 106
- “still” in arguments, 52
- stipulative definitions, 301
- Stradonitz, Kekule von, 312
- Strauder v. West Virginia*, 362
- straw men, 52, 64–65, 341–343
- strength, as inductive standard, 180–182
- strict scrutiny test, 363–364, 381, 382
- Studies in the Way of Words* (Grice), 39n6
- subarguments, arrangement of, 85–90, 102
- subject, changing, 32
- subject term, 164, 167
- substitution instance, 115–116
- sufficient condition test (SCT), 225–226, 228
- application of, 231–232
  - causal reasoning, 217–221
  - as rigorous test, 234–235
- sufficient conditions, 217–219
- suppressed premises, 96–104. *See also* premises
- argument reconstruction, 102–103
  - and circularity, 324
  - contingent facts, 97–99
  - enthymeme, 97
  - evaluative, 100–101
  - reconstruction method, 102–103
  - uses and abuses of, 100–102
- suspect classification, 363
- syllogisms, 166–175
- definition, 167
  - disjunctive, 125
  - probabilistic version of, 193n1
  - problems in applying theory of, 173–175
  - testing for validity, 169–173
  - Venn diagrams for, 168
- symbols
- ampersand, 115
  - horseshoe, 137
  - therefore, 7–8, 46
  - tilde, 123
  - wedge, 122
- syntactic ambiguity, 293
- syntactic conventions, 21
- syntactic rules, 20
- System of Logic*, A (Mill), 328
- systematic definitions, 302–303
- tangents, 79–80
- target feature, 220

- term clarification, in deep analysis, 83  
terrorist attacks, 320  
theists, 449  
“then” in arguments, 42  
theoretical definitions, 302–303  
thereby test, 23–24  
“therefore” in arguments, 41–42, 46  
“therefore” symbol, 7–8, 46  
Thomson, Judith Jarvis, 392–394  
“though” in arguments, 51–52  
“Threat Down, The” (Colbert), 320  
*Through the Looking Glass* (Carroll), 18  
tilde, 123  
torture, 305  
tradition, appeals to, 319  
*Treatise of Human Nature* (Hume), 193  
“trick” argument, 66  
trick of abusive assurances, 49  
trick of excess verbiage, 80  
trick of the disappearing guard, 51  
“Trivia and Significa” (Smith), 109–110  
true positives, 253–254  
Truman, Harry, 186–187  
truth, 19, 32, 93–94, 102  
truth conditions, 114–115  
truth function, 126–128  
truth tables  
    for conditionals, 136–142  
    logical connectives and, 133–134  
    testing for validity, 129–130  
    vs. Venn diagrams, 163  
truth-functional connectives, 126–128  
truth-functionally equivalent, 133–134  
*tu quoque*, 311  
Tversky, Amos, 243, 261n1, 261n3
- uncertainty, decision under, 268  
uncertainty, exploitation of, 390  
unclarity, uses of, 275  
unfairness, 391  
United Nations, 79–80  
*United Steelworkers of America v. Weber*, 368  
universal affirmatives, 155  
universal claims, 334–335
- universal conclusion, 183  
universal discounting, 329–330  
universal negatives, 155  
universal propositions, 155  
universe, origins of, 451–452  
University of California at Davis, 364  
University of Michigan, 368–369  
“unless” in arguments, 148  
unsoundness, 94  
unstated premises, 100  
*U.S. v. Cruikshank*, 362  
*Utilitarianism* (Mill), 296  
utility, expected, 267
- vacuity, fallacies of. *See* fallacies of vacuity  
vagueness, 83, 276–278, 390. *See also* fallacies of vagueness  
vagueness, fallacies of. *See* fallacies of vagueness  
validity, 91–93  
    argument reconstruction, 102–103  
    for categorical arguments, 162–175  
    and circularity, 323–324  
    concept of, 113  
    as deductive standard, 180–182  
    vs. following from, 91  
    Grice on, 144–145  
    in propositional logic, 152  
    vs. soundness, 94  
    testing for, 128–132, 144, 163  
    testing syllogisms for, 169–173  
Vallone, Robert, 261n3  
Venn, John, 153  
Venn diagrams  
    basic categorical forms, 155–156  
    categorical propositions, 152–155  
    contradictories, 159–160  
    definition, 153  
    immediate inferences, 164–166  
    for syllogisms, 168  
    testing syllogisms for validity, 169–173  
    vs. truth tables, 163  
    validity for categorical arguments, 162–175  
viability, as condition of personhood, 388–389

Vietnam War, 285  
violinist analogy, and abortion,  
392–393

“Watch and the Watchmaker,  
The” (Paley), 212–213  
water, 305n5  
wedge, 122  
“whereas” in arguments, 52  
“while” in arguments, 52

White, Tim, 210–212  
Wilberforce, Bishop, 338  
Wolf, Carl, 296  
worldviews, as self-sealing,  
329–330  
worst of crimes argument, 412  
Wright, N.T., 455  
“yet” in arguments, 52  
*Yick Wo v. Hopkins*, 357–360, 362



