4

Reality

According to chapter 3, inference is the process of accepting mental representations on the basis of coherence: we infer a representation if incorporating it with the rest of our representations maximizes coherence. Knowledge, in the sense used by philosophers, requires more than inferring mental representations, which constitute knowledge only if they are representations of reality. But what is reality, and what is fundamentally real? These are the basic questions of metaphysics, and this chapter uses the ideas about coherence developed in the previous two chapters to address important metaphysical questions. I begin with a discussion of the nature of truth and argue *against* a coherence theory of *truth*, as opposed to the coherence theory of *knowledge*, which was defended in chapter 3. I will argue that the coherence theory of truth fails because explanatory coherence supports the existence of a world independent of our thought of it, so that truth must be a matter of correspondence with this world. I then show how coherence as constraint satisfaction provides a new way of thinking about correspondence and approximate truth, by construing modeling the world as a coherence problem.

Explanatory coherence is also the key to fundamental questions about the nature of mind. I defend a materialist view of mind and mental processes that rejects any aspect

of mind concerned with soul or spirit. Inference to the best explanation of mental phenomena does not require any dualist hypotheses that postulate nonmaterial substance. Explanatory coherence combines with analogical coherence to provide the best solution to the philosophical problem of other minds (that is, whether there are any), and it combines with other kinds of coherence to provide the best solution to the psychological problem of how we gain knowledge of other minds. Finally, I present a coherence-based answer to the question of the existence of God. A thorough discussion of this issue would require a book in itself and I do not claim to provide a definitive solution. Rather, my goal is to illustrate how coherence-based inference can be applied to metaphysical issues.

This chapter is intended to be of both philosophical and psychological interest. The metaphysical issues it addresses are some of the most basic in philosophy, but all have interesting psychological analogs. "Is there a world?" may seem like a puerile question suitable only for introductory philosophy classes and adolescent bull sessions, but the psychological and epistemological question "How do we know about the world?" is worthy of adult discussion. Thus in addressing the nature of reality I will simultaneously be discussing the nature of the mental processes that bring us knowledge of the world.

1 TRUTH AND THE WORLD

Philosophical concern with coherence arose with idealist philosophers such as Hegel (1967) and Bradley (1914). Idealism, in the metaphysical sense, is the claim that reality is fundamentally dependent on mind. It contrasts with materialism, which views reality as consisting of matter that is not mind-dependent and provides the basis for

mind, which is viewed as just another function of the physical body. Idealism fits naturally with a coherence theory of truth, according to which a representation such as a proposition is true if and only if it is included in the most complete and maximally coherent set of propositions. On this view, truth just is coherence, since reality is essentially mental and there is nothing outside mind and coherence for a representation to correspond to. In contrast, the correspondence theory of truth, which dates back at least to Aristotle, says that a proposition is true if and only if the world is as the proposition says it is.

For some philosophers, talk of *the world* independent of our minds seems problematic. We have no direct access to this world, and our knowledge of it comes at best indirectly, through sensory experiences and reasoning based on them. What knowledge we may have is inescapably fallible, depending on experiences that may be illusory and reasoning that may be fallacious. All we have, the idealist says, is a complex mix of representations that must be assessed with respect to their coherence with each other, not with respect to some unattainable standard of correspondence to an unreachable and ineffable world.

But the coherence theory of truth has problems of its own. First, there are the questions about isolation from reality and indiscriminate treatment of propositions discussed in the last chapter. I argued that these problems can be overcome only by appreciating some elements as favored, but did not explain why the results of sensory observation and experiments based on it should be favored. Special status does not derive from certainty, for in a coherentist epistemology any observation can potentially be overridden on the grounds that it does not cohere with all the rest that we know. For example, my perception of a giant purple moose on skates should not immediately lead to the inference that there is a purple moose

in front of me, because alternative explanations (such as that I am hallucinating) should be considered for such an unusual observation. In science, it is commonplace for physicists, psychologists, and other experimenters to reject data that they have reason to believe are faulty, for example, because the observations were based on defective instruments or were outliers with respect to other results. Nevertheless, scientists do not treat experimental results as arbitrary and fanciful: some data are thrown out for good reasons.

Observation does not provide guaranteed access to reality, but there are aspects of observation which only make sense if we understand it as being caused by an external reality rather than being a purely mental operation. Here are some aspects of observation that are difficult to explain within a purely coherentist, idealist perspective:

• People cannot observe what they want: most sensory experience is beyond conscious control.

• Different people in the same situation report very similar experiences. For example, just about everyone at a St. Louis Cardinals baseball game will see Mark McGwire hit a home run at the same time.

• Observations of rocks, fossils, and archeological sites suggest that the planet Earth has existed for billions of years, but that humans have existed only for a few million.

Thus human observation is comparatively mind-independent, intersubjective, and historically recent.

Materialism has no difficulty explaining these three facts. First, observation by an individual is largely mind-independent because it is the result of physical processes involving causal interactions between our sense organs and a material world. Second, observation is intersubjective because different individuals share very similar sensory organs and all operate in the same world. Third, the rela-

tive recency of human observation is explained by scientific theories about the development of the universe and the solar system and about the much more recent evolution of the human species.

These aspects of observation make sense within idealism only if there is some kind of collective or divine mind that determines the experiences of diverse people and that has contrived to present the appearance of a world in which people are relatively recent arrivals. But we have no independent reason to believe in such a collective or divine mind (see the discussion of God below), so the idealist explanations are lacking in simplicity. As principle E2c in the first section of chapter 3 specified, lack of simplicity deriving from introduction of additional hypotheses reduces the explanatory coherence of a theory. Leaving aside for now the issue of God, it is clear that the idealist hypothesis that the world is constituted by and dependent on mind has less explanatory coherence than the materialist hypothesis that the world consists fundamentally of molecules, atoms, subatomic particles, quarks, and the other physical entities that science has discovered.

So there is a material world independent of our minds' most coherent interpretations of it. The truth of propositions, therefore, and the verisimilitude of other mental representations, such as visual images, is not merely a matter of their coherence with other representations; rather, the truth of a proposition depends on its correspondence to the world. Of course, we have no other means but coherence to infer that a proposition does correspond to the world, and any particular proposition, no matter how coherent, may turn out to be false. But we have ample reason to believe that many bodies of propositions do in fact correspond to reality. People have managed to survive and reproduce for many thousands of years, and in the last few hundred they have been able to use scientific advances

to gain an extraordinary physical control over the world. These advances have made possible technologies of transportation, communication, and medicine that are totally mysterious unless such scientific theories as gravity, electromagnetism, and the germ theory of disease are at least approximately true.

Socrates, Descartes, and other philosophers who have taken skepticism too seriously have unfortunately set the stage for much of the history of philosophy by asking for definitive proof for philosophical theses. Hegel and Peirce were the first philosophers to recognize that the foundational search for certainty was pointless, and that what mattered was the growth of knowledge, not its foundations. Back in 1868, Peirce urged, "Let us not pretend to doubt in philosophy what we do not doubt in our hearts" (1958, 40). What matters is how we use inferential processes to expand our knowledge and eliminate previous misconceptions. Peirce did not recognize, as Hegel did obscurely, that building inferences upon each other is based on expanding sets of coherent representations. However, Peirce's notion of abductive inference to explanatory hypotheses is the ancestor of contemporary ideas about inference to the best explanation and explanatory coherence.

2 CORRESPONDENCE AND APPROXIMATE TRUTH

But what is it for representations to correspond to the world? Philosophers of science have identified serious problems with the realist claim that scientific theories are true. Cartwright in *How the Laws of Physics Lie* (1983) considers physical laws as idealized rather than exactly true descriptions of the world. For example, Newton's law of gravitation, that two bodies exert a force between each

other that varies inversely as the square of the distance between them and varies directly as the product of their masses, is not true of bodies that are electrically charged. Giere (1999) argues that the point of scientific theories is not directly to make claims about the world, but rather to define *models* that fit the world more or less well. A model, on his view, is a nonlinguistic entity that has the same relation to the world that a map has to the aspects of the world that it is intended to represent. Maps are not absolutely true or false, but can be more or less accurate and more or less detailed.

The relation between models and reality is not isomorphism, which would require there to be a one-to-one mapping between the model and the world that exactly preserves structure and behavior. Holland et al. (1986) describe the relation between a model and the world with the ugly term *quasi-homomorphism*, which means a mapping from the model to some parts of the world in which the model preserves many but not all of the structures and behaviors in the world. Assessing the relation between a model and what it represents can be viewed as a coherence problem very much like the process of analogical mapping described in chapter 3. Analogies are rarely isomorphic to each other, but can nevertheless involve correspondences between two analogs that are highly useful. Similarly, models, like maps, can provide more or less coherent representations of the world.

To show this more exactly, we need more concrete specifications of what models are. As Giere insists, theories and models are representationally heterogeneous, involving diagrams and pictures as well as equations and other propositions. So the mapping by which a theory defines a model can be complex, and the models can be entities that bear visual as well as semantic correspondences to the world. For simplicity, we can begin with the

set-theoretic notion of model used in Tarskian semantics for formal languages. On this usage, a model consists of a domain D_M, which is a set of objects, and a collection R_M of relations, which are n-tuples of objects in D_M. For example, a simple domain consists of the domain {Bill, Tony, Phil} and the relation {(Bill, Tony), (Bill, Phil)} which might be interpreted as saying that Bill is taller than Tony and Phil. Properties are construed as one-place relations, corresponding to sets of the form {(o), ...}. Similarly, we can say that the world consists of a domain of objects D_W and a set of relations R_W among the objects. The model represents the world to the extent that there is a mapping from D_M to D_W and from R_M to R_W such that the relations in R_M have corresponding relations in R_W.

Because isomorphism is too much to expect of an approximating model, we can think of the world as providing constraints on the model, without the expectation that all the constraints will be satisfied. The degree of coherence between the model and the world is then measured by the degree of constraint satisfaction, W/W*, as defined at the end of chapter 2. In this coherence problem, the elements are set-theoretic relations, such as (Bill, Tony) is a member of {(Bill, Tony), (Bill, Phil)}, and the constraints are between model elements and world elements: a model element is to be accepted if and only if the corresponding world element is accepted. The coherence of the model with respect to the world is then measured by W/W*, the ratio of total weight of constraints satisfied to the total weight of all constraints established by the mapping between the model and the world.

Coherence can similarly be defined using more interesting kinds of models than the Tarskian kind. For example, in physics and other mathematical fields, a dynamic system is often conceived in terms of a mathematical space with orthogonal coordinate directions rep-

resenting each of the variables needed to specify the instantaneous state of the system (Baker and Gollub 1995). The state space of a system is the set of states it can be in as determined by the variables that are used to measure it. For example, the state of a particle moving in one dimension is specified by its position and velocity. Relations between variables are specified by equations that define a trajectory of a particle through the state space. We can think of the equations as defining a model state space that is intended to correspond to the state space of the actual world. For example, an equation with three variables generates a three-dimensional state space. A particular state of the world can be specified by a list of numbers called a vector, which contains the values of all the variables at a particular time. The model as defined by the equations specifies possible transitions from one vector to another, so that changes in the system can be specified by lists of vectors or, equivalently, by diagrams that draw a picture of the relations between vectors in spaces of two or more dimensions. In this framework, an element is a representation of the relations between vectors in the state space. The elements involving the trajectory in the model's state space, S_M, are constrained by the elements involving the trajectory in the world's state space, S_W, which contains the actually occurring sequences of vectors (values of variables). As in the Tarskian case, a model is coherent to the extent to which it satisfies constraints directing the co-acceptance of elements from the model and the corresponding elements from the world.

Different kinds of models will require different kinds of elements, but it should always be possible to define the coherence of the model in terms of the extent to which it satisfies constraints between elements of the model and elements of the world. Hence the notion of fit between a model and the world can go beyond Giere's map analogy

and be specified more generally as a coherence problem. While metaphysically useful, this view of correspondence between a model and the world is of limited practical use. Many aspects of the world that models are intended to capture are not directly observable: we can not directly measure the properties and relations of hypothetical entities such as subatomic particles. Hence practically, we can at most assess the degree of the fit between the elements of the model and world that concern observable objects, and use that to guess the overall fit between all the elements of the model and the corresponding elements of the world. Scientists can proceed more directly, by addressing the question of whether one theory provides a better explanation of the evidence than its competitors, in accord with the theory of explanatory coherence. If one theory coheres with the evidence better than other available theories, then we have reason to believe that it models the world more coherently than the alternatives.

3 MIND AND BODY

Aside from a few wild-eyed philosophers, everyone agrees that there is a world apart from the minds that contemplate it. But no such universal consensus exists concerning the nature of mind. Most cognitive scientists, including researchers in neuroscience, cognitive psychology, and philosophy of mind, adopt a materialist perspective according to which all aspects of mind are ultimately explicable in terms of the brain and the body and world that it inhabits. In contrast, most ordinary people are dualists, taking it for granted that a person is not just a body, but also consists of a nonmaterial soul or spirit. Religions such as Christianity that assume survival beyond death are the source of the everyday metaphysical beliefs that portray a

person as a combination of body and soul. What is the evidence for dualism and materialism, and which metaphysical theory is more coherent?

Support for materialism has increased dramatically in recent decades, as neurological data has made possible detailed explanations of more and more mental phenomena in terms of the operation of the brain. Scientists can measure the firing of single cells in brains of cats and monkeys, and correlate these occurrences with the animals' mental activities. Such invasive measurements are not permissible in humans, but other techniques such as brain scans using positron-emission tomography and functional magnetic-resonance imagery have begun to provide detailed information about what the mind is doing during diverse mental operations, including visual perception, word recognition, and memory. In addition, much has been learned about the operation of neural networks in the brain, including how neurons transmit electrical signals to each other and adjust their connections with each other on the basis of experience. Such progress supports the claim that mind can be understood on the basis of the principles of physics, chemistry, and biology that we use to explain the material world in general.

The metaphysical hypothesis of materialism coheres with scientific findings by providing a higher-level explanation of why the operations of the mind and the brain are so strongly correlated. Figure 4.1 shows in simplified form the explanatory structure of materialism. The top-level metaphysical hypothesis says that everything is matter, i.e., that whatever exists consists of physical entities such as quarks, electrons, atoms, molecules, cells, and organisms. Materialism about mind follows deductively from this universal materialism, which of course must be rejected if the mind is not scientifically explicable. But figure 4.1 also displays the explanatory coherence of materialism about the
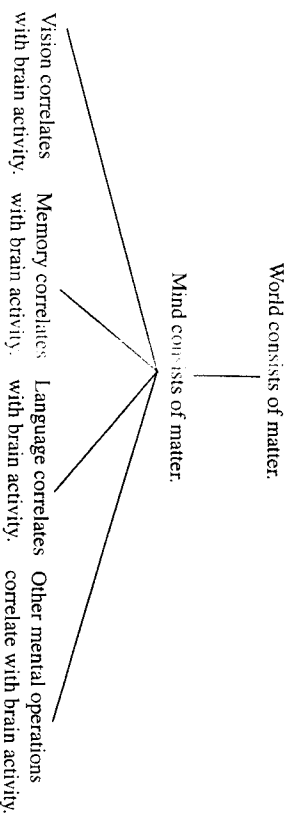
World consists of matter.

Mind consists of matter.

Vision correlates with brain activity. Memory correlates with brain activity. Language correlates with brain activity. Other mental operations correlate with brain activity.

Figure 4.1
The coherence of materialism. The lines indicate positive constraints based on explanation. (See the appendix to this chapter for a much fuller exposition.)

mind, which explains why neuroscience has found such striking correlations between a broad range of mental operations and increasingly well-identified operations in the brain.

Materialism, however, is not the only general way of explaining mental operations. According to idealism, minds operate because everything is mind, but I argued in the last section that idealism is implausible. More plausible is the dualist view that allows that *some* aspects of mind are indeed explicable in material terms, because a person is indeed a body as well as a soul. But the dualist claims that there are other aspects of mind that require explanations based on the existence of a nonmaterial soul. Materialism allegedly is unable to explain such apparent phenomena as the survival of life after death, extrasensory perception, free will, the moral sense, and consciousness.

Let us construct the most powerful case we can for dualism's having greater explanatory coherence than materialism. It would seem that dualism can explain everything that materialism does by virtue of a person having a body, but in addition there is a lot that dualism can explain

but that materialism cannot. Dualism is capable of explaining the existence of life beyond death, although survival is itself a hypothesis whose explanatory coherence must be evaluated. The ability of the soul to survive the body would explain various alleged occurrences, such as near-death experiences when people report going through a tunnel toward a bright light and seances when people report communicating with the dead.

Another phenomenon that would seem to require a dualist explanation is extrasensory perception. Suppose, as some advocates of ESP claim, that people are capable of remote viewing, seeing scenes thousands of miles away without any mechanical devices. And suppose that people can perform telepathy, transmitting information from mind to mind without any physical means of transmission. And suppose that people are capable of telekinesis, affecting matter without any physical contact or connection. These kinds of ESP would provide strong support for dualism, because their existence and operation violate current scientific theories and therefore cannot be explained by them.

Other phenomena that have been given a nonmaterialist explanation include the widespread beliefs in free will and the moral sense, or conscience. We certainly feel that we are acting freely and that we have a sense of right and wrong, and it is not easy to imagine how these feelings can emanate from material processes involving neurons and brain chemicals. The existence of a nonmaterial soul would explain why we appear to have free will and why we can make intuitive moral judgments.

Our experiences of acting freely and making judgments of right and wrong are part of a more general aspect of mind, consciousness. We not only think; we are aware of our thinking, especially of visual and other sensory experiences, as well as emotions and moods.

Consciousness does not seem to be a process like the physical, chemical, and biological ones found in materialist explanations. The dualist contends that consciousness needs a different kind of explanation: only nonmaterial souls are capable of the awareness and qualitative experiences that constitute consciousness. Like survival after death, ESP, free will, and moral intuition, consciousness requires explanation by a nonmaterial component of mind. Figure 4.2 sketches the explanatory coherence of dualism.

The materialist, however, has a good shot at explaining all these phenomena. Materialism cannot explain how minds could survive without brains, but it can explain why people report communication with the dead and near-death experiences. Seances are easily staged, so the materialist can explain them as fraudulent performances. Near-death experiences can be explained neurologically and socially: It is possible that the process of expiring produces a flood of brain chemicals such as endorphins that generate the unusual experiences reported by people who have come close to death. The similarity of the reports may be explained by people having similar brain chemistry, but also by people near death having previously heard of the experiences of other people.

Reports of extrasensory perception can also be explained in materialist terms. Many attempts to demonstrate the occurrence of anomalous phenomena such as remote viewing, telepathy, and telekinesis have been exposed as fraudulent or inadequately designed to rule out chance or bias as alternative explanations of the alleged results. A few attempts have been made to determine the existence of ESP with full scientific rigor, but at best the effects found have been very small and explicable by the alternative hypotheses of fraud or poor experimental design. Because materialism provides explanations of ESP
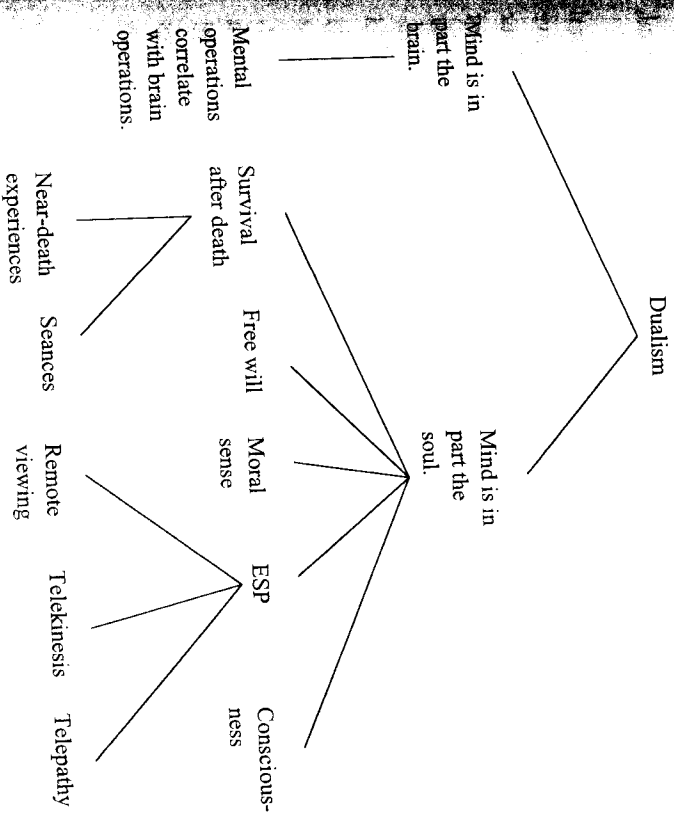

Figure 4.2
The explanatory coherence of dualism. Lines indicate explanatory relations.

that are at least as plausible as the existence of a soul, dualism gains little support from ESP.

Similarly, the semblance of free will and the moral sense can be explained in material terms. We think we have free will because we are not aware of the operations of our brains, just as prescientific people explain the operation of their animals and the weather in terms of spirits because of their ignorance of the underlying physical processes. The illusion of free will derives partly from the religious and dualistic culture that most of us are raised in, and partly from the reality of conscious experience that I will discuss shortly. Although we do not have free will in any absolute, dualist sense, we certainly are capable of acting freely in a social sense when our choices are not directly controlled by others or by the defective neurochemistry that produces schizophrenia and other mental disorders. Such freedom suffices for moral and social responsibility: we hold people responsible not because they are free in the sense required for religious doctrines of sin and divine punishment, but because doing so helps to produce desirable outcomes for them and for other members of society.

From a materialist perspective, the existence of moral intuition can be explained in the same way that other kinds of intuitive judgments are explained. When we compile various kinds of information into an overall judgment of the rightness or wrongness of some act such as murder, we are unconsciously producing a coherence judgment that can have a strong emotional content. Chapter 5 describes how ethical judgments can arise from a combination of coherence-based inferences, and chapter 6 shows how these can be combined with emotional evaluation.

Materialism has a fairly easy time offering explanations of near-death experiences, ESP, free will, and moral judgment, but the general phenomenon of consciousness is much more difficult. Even if we find all sorts of neural

correlates of sensory experience and awareness, the dualist can still maintain that neurology has not shown how consciousness is produced by the brain. Recently, however, brain scientists have begun to develop detailed hypotheses about the neural origins of some aspects of consciousness. For example, Crick (1994) uses what is known about the brain's visual, attentional, and memory systems to conjecture how consciousness might emerge from circuits of interacting neurons. Thus the materialist explanation of consciousness, that it emerges from neural processes, is sketchy but promising.

Some philosophers, e.g., Chalmers (1996), think that consciousness can be used to support dualism on purely conceptual grounds. For example, we can conceive of the existence of zombies, which are physically identical to humans but lack consciousness, so that consciousness is logically independent of bodies. But conceivability is a poor guide to reality, and dualism must be evaluated with respect to its explanatory coherence, not on conceptual grounds tainted by prior beliefs.

To fully assess the competing explanatory coherence of materialism and dualism, we need to combine all the dualist and materialist explanations shown in figures 4.1 and 4.2. The outcome is a much closer call than the debate between idealism and materialism, for there are aspects of mind such as consciousness that are not obviously within the scope of scientific explanation. But even if the materialist explanation of consciousness is currently weak, so is the dualist explanation, for no one has offered any account of how the soul produces consciousness that is any more successful than one for the brain. In the appendix to this chapter I present a fuller analysis of the explanatory coherence of dualism and materialism, which needs to be connected with the question of the existence of God. Dualism is much more plausible if God created souls, and theism

loses plausibility if materialist explanations of the existence of the universe are available. Hence the mind-body problem involves theological issues, which are addressed below. First I want to address another traditional philosophical question: whether there are other minds.

4 OTHER MINDS

Here is the traditional philosophical problem of other minds: you know from your conscious experience that you have a mind, but how are you justified in believing that other people, whose consciousness you have no access to, have minds? Like the problem of whether there is a mind-independent world, this problem is rather silly, since no one doubts that there are other minds. We can dispose of the philosophical problem quickly, then move on to the more interesting and pressing psychological question: given that there are other minds, what can we know about them?

One common solution to the philosophical problem is analogical inference: other people's actions are similar to yours, so perhaps they are also similar to you in having minds. Another common solution to the problem of other minds is inference to the best explanation: the hypothesis that other people have minds is a better explanation of their behavior than any other available hypothesis, for example, that they are radio-controlled robots. From the perspective of coherence as constraint satisfaction, analogical inference and best-explanation inference are complementary, not alternative justifications, because analogical- and explanatory-coherence considerations can simultaneously work to justify as acceptable the conclusion that other people have minds. Figure 4-3 shows how analogy-based positive constraints mesh with
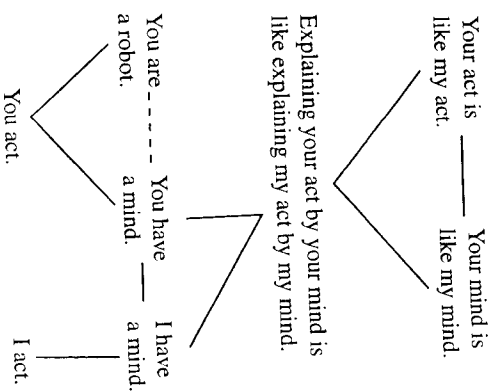
Your act is ——— Your mind is
like my act.   like my mind.

Explaining your act by your mind is
like explaining my act by my mind.

You are - - - - You have   I have
a robot.   a mind. —— a mind.

     You act.   I act.

Figure 4-3
Support for the existence of other minds incorporating both explanatory and analogical coherence. Solid lines indicate positive constraints and the dashed line indicates a negative constraint. My hypothesis that you have a mind is evaluated both by comparing its explanatory power with other hypotheses that explain your behavior and by analogy with my explanations of my own behavior.

explanation-based positive constraints to establish the acceptability of the hypothesis that other people have minds. The hypothesis that another person has a mind is supported both by its greater explanatory coherence over competing explanations of your actions and by its analogical coherence based on the similarities between your acts and my acts.

So other minds exist, but how do we know them? First, we understand other people by means of causal attributions in which we form and evaluate hypotheses that explain their behavior. To explain why someone is abrupt on one occasion, you may hypothesize that this person is

impatient or that he or she is under pressure from a work deadline. You believe the hypothesis that provides the best available explanation of the person's behavior. A second means of making sense of people is analogy: you can understand people through their similarity to other people or to yourself. For example, you may understand the stresses that your friend is experiencing by remembering an occasion when you yourself experienced similar stresses. This will allow you to predict your friend's likely feelings and behavior.

Causal attribution of mental states is naturally understood in terms of explanatory coherence. The elements are propositions, including the evidence to be explained (observed behavior) and hypotheses about them that would explain the behavior. Suppose, for example, that a normally mild-mannered friend screams at you. Various hypotheses would explain that behavior: perhaps the friend had a stressful day at work, or stopped taking some needed medication, or learned some secret ugly fact about you. What inference you make to explain your friend's behavior will depend on what best fits with your other beliefs: maximizing coherence will lead you to accept the most plausible hypothesis that explains your friend's behavior and to reject the alternative hypotheses. You may hypothesize that your friend screamed at you because of a stressful day at work and further hypothesize that the stressful day was caused by impending layoffs. The result can be a network of propositions of the sort shown in figure 4.4, which shows different hypotheses competing to explain the evidence. Positive constraints can be affected by considerations of simplicity: if a number of hypotheses are required to make an explanation, then the positive constraints between hypotheses and evidence are weakened. For example, if you explain Mary's behavior by supposing that she was
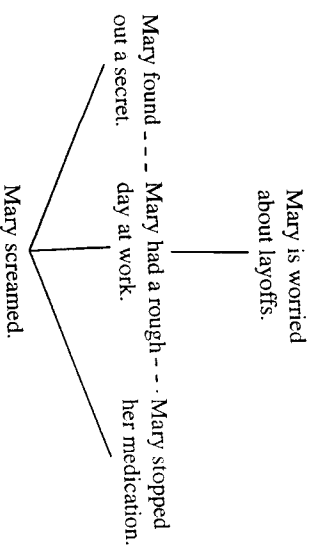
Mary is worried
about layoffs.

Mary found - - - Mary had a rough - - - Mary stopped
out a secret.   day at work.   her medication.

Mary screamed.

Figure 4.4
An explanatory-coherence network for Mary's screaming. Positive associations are shown with solid lines and negative associations are shown with dashed lines. The evidence that Mary screamed can be explained by three competing hypotheses.

abducted by aliens who mistreated her, you are making a number of hypotheses whose coherence may suffer as a result of a lack of simplicity as well as incompatibility with other things that you believe.

In explanatory coherence, the sources of negative constraints are contradiction and competition. If two propositions logically contradict each other ("Mary is in Florida" versus "Mary is in Toronto"), then there is a strong negative constraint between them. Moreover, in explanatory situations, people tend to treat hypotheses as negatively constraining each other even if they are not strictly contradictory. It is possible that Mary's behavior should be explained because she had a stressful day *and* she stopped taking her medication *and* she found out something about you, but in the absence of evidence linking them, we treat these as independent competing explanations. Explanatory coherence can also be used to assess hypotheses about oneself, as when Mary herself figures out that she screamed because of some previously suppressed hostility.

Another valuable cognitive mechanism for making sense of people is analogy, in which we see one person as similar to another with respect to a complex of properties and relations. I may, for example, increase my understanding of Princess Diana by comparing her to Anna Karenina in Tolstoy's novel. This comparison would be much deeper than just noticing that both are women who came to tragic ends, in that it also involves a set of interlocking relations. Diana was like Anna Karenina in being married to a man, not caring for that man, and being (for a while) passionately involved with another man. The analogy involves noticing not only that Diana corresponds to Anna, but also that Prince Charles corresponds to Anna's husband, and that Diana's lover James Hewitt corresponds to Anna's lover Vronsky.

As we saw in chapter 3, such analogical mapping can be viewed as a coherence process that maximizes the satisfaction of multiple constraints. The elements are hypotheses about what corresponds to what, for example, that Diana corresponds to Anna and that *loves* in Diana's case corresponds to *loves* in Anna's case. One constraint is perceptual and semantic similarity: two elements will tend to correspond to each other if they look the same or have similar meaning. Other constraints are structural: to map *Anna loves Vronsky* to *Diana loves James*, we must consistently map Ann to Diana, loves to loves, and Vronsky to James. Mappings should tend to be one-to-one; without strong reason, we should not map Diana to both Anna and Vronsky. Finally, purpose provides a practical constraint on the mapping, because we should try to come up with mappings that will contribute to the cognitive goals that the analogy is supposed to serve, such as providing an explanation or contributing to a decision.

Analogically making sense of people always involves comparing two individuals, a target to be understood and

a source that provides understanding. In the Princess Diana example, the source and target are both other people, but sometimes the source is oneself and the target is another (e.g., empathy), sometimes the target is oneself and the source is another (e.g., some kinds of social comparison), and sometimes both the source and target are oneself (as when a past situation of one's life is used to make sense of a current situation). I can get a better understanding of my own mind by comparing my current situation with a previous one where I knew what I was thinking and feeling.

Empathy is analogical mapping from another to oneself, establishing a correspondence not only between someone else's situation and one's own but also between the other's emotional state and one's own emotional experience. Deep understanding of people's work stress requires not just seeing how their situation corresponds to one that I have been in (unpleasant boss, risk of layoff, etc.) but also appreciating their emotional state (anger, fear). In a purely verbal analogy, I may infer that just as I was angry in my own situation, so the other is likely to be angry in a similar situation. But empathy goes beyond verbal elements by providing a correspondence between some emotional experience of my own and what I can analogically infer to be the emotional experience of the other. By setting up an analogy between another person and myself, I can feel an approximation to what the other feels. Such an analogy should be facilitated if I myself have been in a similar situation. Indeed, Batson et al. (1996) found that women felt greater empathy for someone undergoing a difficult experience if they themselves had had a similar experience (though the same was not true for men). Like other kinds of analogy, empathy can be understood as a coherence mechanism that evaluates a set of correspondences between two people and their situations; empathy differs

from other analogies in that the correspondences link representations that are not verbal or visual, but emotional. Additional discussion of empathy must await the account of emotional coherence presented in chapter 6.

We take for granted our ability to understand, at least to a large extent, the minds of others. But people with autism have a greatly diminished ability to comprehend other minds. This deficit has led some autism researchers to postulate that there is an innate mental module for a "theory of mind" that enables people to understand each other. But Uta Frith (1989) cites many studies that show that autistics' problems are much more general than the inability to understand other minds. Their other defects in visual and linguistic reasoning suggest that autistic people suffer from the more general defect that Frith calls "weak central coherence." O'Laughlin and Thagard (forthcoming) have used coherence models to simulate how autistics make defective inferences in both theory-of-mind and language-understanding tasks. Five-year-old children, unlike three-year-old children and older autistics, can make powerful inferences about the mind of another child who does not see a marble moved from one place to another. Older children infer that the other child will look for the marble in the place that the other thinks it is, rather than in the place where it really is. This sophisticated kind of explanatory-coherence-based inference breaks down when the coherence model is distorted by making inhibition very strong in comparison with excitation, resulting in preference for the most immediately appealing conclusion rather than in the conclusion that maximizes coherence. Similarly, autistics have difficulty processing sentences such as "The girls were climbing over the hedge. Mary's dress remained spotless, but in Lucy's dress there was a big *tear*" (Frith and Snowling 1986). Here "tear" needs to be interpreted as a hole in the dress, not as water in the eye, but autistics

are not able to use the context of the sentence to make the interpretation that best fits overall. O'Laughlin and Thagard (forthcoming) show that the same coherence defect (an excess of inhibition over excitation) that produces incorrect inferences in the connectionist coherence model of children's false beliefs also reproduces the autistics' incorrect inferences in the dress example. Thus the theory of coherence as constraint satisfaction has the potential to explain some failures to understand other minds as well as numerous successes.

5 GOD

We can also use the theory of coherence to address another major metaphysical question: does God exist? I shall produce what I think is the best possible argument for the existence of God, based on explanatory and analogical coherence. It will turn out, however, that a full assessment of the coherence of theism, which asserts the existence of an all-powerful nonmaterial being and thus contradicts materialism, supports the conclusion that there is no reason to believe in the existence of a divine being.

For many religious people, belief in the existence of God is not a matter of evidence or reason, but of faith and tradition. But some theists have defended the explanatory coherence of the existence of God. Swinburne writes,

The basic structure of my argument is this. Scientists, historians, and detectives observe data and proceed thence to some theory about what best explains the occurrence of these data. We can analyse the criteria which they use in reaching a conclusion that a certain theory is better supported by the data than a different theory—that is, is more likely, on the basis of those data, to be true. Using those same criteria, we find that the view that there is a God explains everything we observe,

not just some narrow range of data. It explains the fact that there is a universe at all, that scientific laws operate within it, that it contains conscious animals and humans with very complex intricately organized bodies, that we have abundant opportunities for developing ourselves and the world, as well as the more particular data that humans report miracles and have religious experiences. In so far as scientific causes and laws explain some of these things (and in part they do), these very causes and laws need explaining, and God's action explains them. The very same criteria which scientists use to reach their own theories lead us to move beyond those theories to a creator God who sustains *everything* in existence. (1996, 2)

Thus according to Swinburne, belief in the existence of God does not require a leap of faith, but can arise from the same explanatory reasoning found in science.

At first glance, the hypothesis that there is a God does seem to have a great deal of explanatory coherence. First, it explains why the universe exists, i.e., because God created it. One traditional argument for the existence of God contends that everything has a cause, so the universe must have a cause, namely God. This is not a deductive argument, for it does not show that the cause of the universe is the omnipotent being that theists usually take God to be. Rather the cosmological argument, as it is called, is best construed as an inference to the best explanation: we should accept the hypothesis that there is a God because it provides the best explanation of the existence of the universe.

But the existence of God can explain more than the universe's existence: it can explain why the universe is as it is, with the specific scientific laws that govern it. Physical laws such as Newton's laws of motion and biological laws such as genetic transmission hold because God designed them that way. Design is especially important for explaining the complexity of biological organisms such as

humans. The traditional argument from design says that God is responsible for the wonderful abilities of organisms to function in the world. The argument is partly a matter of explanatory coherence: God's plan explains the complexity and adaptations of organisms. But it also involves an analogy between God's design and human design. William Paley (1963) compared the complexity of the world to that of a watch and argued analogically that just as a watch has a designer, so does the physical and biological world. Intricately adapted organs such as the eye are taken as signs of God's existence.

If, contrary to the argument made earlier in this chapter, humans consist of souls as well as bodies, then God's existence can be used to explain the existence of souls. Because souls are nonmaterial, their existence is not explicable scientifically, so a different metaphysical explanation is required. Souls exist because an all-powerful nonmaterial being created them.

The existence of God would also provide an explanation for miracles and religious experience. Miracles occur because God occasionally intervenes in the world, and he sometimes interacts with people, providing them with religious experiences. In addition, many people believe that God is the source of morality, providing an explanation of why there is right and wrong and why most people believe there is right and wrong. In sum, we get an impressive picture of the coherence of theism, shown in figure 4-5. The figure is incomplete in that it does not show the analogical connections between the explanation of biological complexity in terms of God's design and the explanation of complexity in artifacts in terms of human design.

The explanatory coherence of theism appears overwhelmingly impressive until one begins to examine alternative explanations of the phenomena taken to support it. The metaphysical hypothesis of materialism contradicts
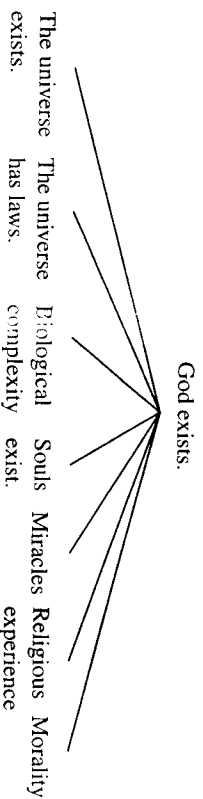
God exists.

The universe exists. | The universe has laws. | Biological complexity exist. | Souls exist. | Miracles | Religious experience | Morality

Figure 4-5
The explanatory coherence of theism.

theism and provides alternative explanations of the apparent support for the hypothesis of God's existence. There are various nontheistic explanations of the existence of the universe: perhaps the universe has always existed, or perhaps it came into being spontaneously as the result of random energy fluctuations in quantum fields. Obviously, we lack good evidence for either of these two hypotheses, but then there is no direct evidence of divine creation either. It might seem that they complicate the materialist hypothesis by requiring extra assumptions in the explanation of the existence of the universe, but the theistic hypothesis also requires additional assumptions, for example, that God decided to create the universe and had the power to do so. With respect to explanation of the existence of the universe, the theistic explanation has no clear advantages over the materialist ones.

Why is the universe governed by its physical laws? The materialist explanation here is partly reductive and partly historical. The reductive part comes from the assumptions that biological laws derive from chemical laws and chemical laws from physical laws, and that physical laws derive from the fundamental forces and particles that operate universally. Although science is not currently able to fill out these derivations completely, there is abundant knowledge of some of the crucial dependencies. For example, biolog-

ical laws of genetic inheritance have their basis in chemical laws involving molecules such as DNA, and chemical molecular interactions are based on the operations of atoms and subatomic particles. Why do these fundamental entities behave the way they do? Here we can at most hazard a historical explanation, based on the early development of matter after the big bang around twenty billion years ago. Science cannot explain exactly why the present laws of nature came to be, but then theism cannot explain why God chose to construct a world that falls under Newton's laws of motion.

Much more can be said concerning the materialist explanation of biological complexity. The argument from design lost its cogency in 1859, when Charles Darwin published *On the Origin of Species*, describing how evolution by natural selection could produce new species with complex organs. Darwin explicitly considered divine creation as the alternative to evolution in explaining biological facts, but mounted a long and impressive argument that organisms including humans could evolve by natural means (Thagard 1992b, chap. 6). In the 140 years since the *Origin* appeared, an astonishing amount of evidence has accumulated that is best explained by the theory of evolution by natural selection, supplemented by the more recent theory of genetics, which explains how traits are transmitted from one generation to another and how variation can occur. The analogy between human design of artifacts and biological design of organisms has been progressively undercut by the substantial amount of evidence that biological complexity can arise from nonintentional means such as genetic variation and natural selection.

A modern version of the argument from design is based on the *anthropic principle*, according to which "all the seemingly arbitrary and unrelated constants in physics

have one strange thing in common—these are precisely the values you need if you want to have a universe capable of producing life" (Glynn 1997, 22). Physicists have argued that if the physical constants such as the values for gravitational force or electromagnetic force had varied much from the actual values, then the universe would be very different and life would not have evolved. Glynn takes the anthropic principle as pointing to a religious explanation of why the fundamental constants have the values that they do. God must have picked those values of the constants in order to ensure that life would evolve. This explanation requires many assumptions: that God exists, that God wanted life like that found on Earth to evolve, and that the only way God could produce such life was by choosing the currently observed values for the physical constants. At best, this is a weak explanation: according to explanatory coherence principle E2c, the more hypotheses it takes to explain something, the lower the degree of coherence. Science does not do any better in explaining the values of the physical constants: there are strange speculations about our universe being one of many evolving from black holes in previous universes. In general, however, the hypothesis that the values of the physical constants are happy accidents is as plausible as the hypothesis that they are the result of divine design, unless one already believes that God designed the world.

The other alleged evidence for God's existence fares even worse in the explanatory battle with materialism. I argued earlier in this chapter that explanatory coherence supports materialism over dualism. Because there is no good reason to believe that souls exist, we cannot use their existence as evidence for the existence of God. Here materialism does not offer an alternative explanation, only a rejection of the alleged fact to be explained. Similarly, miracles are not a fact to be explained: the materialist denies that

they occur. What does need to be explained is that some people report that miracles have occurred, but it is easy to account for these reports on the basis of individual and social delusions. Similarly, psychological explanations are available for why people have religious experiences, which derive from social experiences and individual needs to believe in contact with God. I have already argued, in discussing dualism, that materialist explanations of the moral sense are possible and plausible, so no theistic explanation is needed.

From a psychological perspective, it is misleading to discuss arguments about dualism and arguments about the existence of God separately from each other. Theism and dualism go hand in hand, and evidence for one supports the other. It is difficult to diagram the full complexity of such coherence-based inferences, but in the appendix I provide an encoding of the propositions and coherence relations involved in assessing dualism and theism together as challenges to materialism, and I describe a computer simulation that supports my claim that materialism is more coherent.

So far my analysis of the coherence of theism has concentrated on what the existence of God might be able to explain and has ignored a great deal of evidence that traditional theism has difficulty explaining. The billions of people who have existed during the past one hundred thousand years or so have undeniably undergone a great deal of suffering, arising from famine, war, disease, death, and other afflictions. These occurrences would not be a problem for a theist who believed that a malevolent god created humans in order to observe their pain, but Christians and most other theists believe that God is inherently good and wants the best for people. Thus theism seems to be incoherent with the huge amount of evil in the world. The standard theistic explanation of evil in the world is

free will: it maintains that God wanted people to be free to make their own choices. But, as the discussion of dualism showed, there is reason to believe that free will in the absolute sense is an illusion. And even the assumption of free will does not explain the existence of so much suffering not derived from human actions, such as the occurrence of diseases that cause physical and emotional suffering.

In contrast, there is an obvious materialist explanation of human suffering. People are biological organisms subject to disease, famine, and death just like all other species of animals. We differ from other animals in having greater intellectual capacity, which unfortunately is sometimes used to inflict suffering on other people through wars and other actions. Human suffering thus has natural biological, psychological, and sociological explanations that do not require invoking any extra ill-supported hypotheses such as free will.

As the appendix shows in more detail, the conflict between materialism and theism requires that the latter be rejected as part of the maximally coherent explanation. Why, then, is belief in God so widespread? The reasons are partly sociological, in that people are brought up by parents and other teachers who pass on their religious beliefs. The reasons are also party psychological, in that belief in God provides solace and hope to many people, who otherwise would experience despair at the difficulties that life carries with it. Chapter 6 describes how our coherence judgments are intermixed with emotions.

In his fullest argument for the existence of God, Swinburne (1990) concludes that the hypothesis of God's existence is *probable*, given evidence such as the existence and nature of the world. He neglects to consider that the meaning of probability as applied to explanatory hypotheses is problematic, and that the assessment of

hypotheses requires consideration of alternative explanations. Chapter 8 provides a systematic comparison of the relation of probabilistic reasoning and explanatory coherence.

In the middle of the twentieth century, metaphysics fell into disrepute when the logical positivists contended that metaphysical questions are unanswerable and meaningless because they are not subject to empirical confirmation and refutation. As it turned out, the positivists' view of scientific inference was much too narrow and would have condemned most of science as unscientific. But inference to scientific theories is naturally construed in terms of explanatory coherence (Thagard 1992b), and exactly the same kind of inference can be used to address metaphysical questions such as the existence of God. Analogical and deductive coherence are also relevant. There is, therefore, nothing inherently disreputable about metaphysics, although many who have claimed to pronounce upon the fundamental nature of reality have produced implausible theories. There is no conflict between science and metaphysics, only between science and bad metaphysics. At the edges of science, metaphysical questions about the fundamental nature of reality inevitably arise, and they can be answered by the same kinds of coherence-based inferences found within science.

6 SUMMARY

A coherence theory of knowledge and inference can be used to justify a realist theory of truth, the world, and other minds. Simultaneously, coherence considerations lead one to reject as implausible such nonmaterial entities as spirits, souls, and gods. Just like scientific theories, metaphysical hypotheses about the fundamental nature of

reality can be evaluated with respect to their explanatory and other kinds of coherence.

7 APPENDIX: THE COMPARATIVE COHERENCE OF MATERIALISM, DUALISM, AND THEISM

A full comparative analysis of the coherence of materialism needs to integrate all the hypotheses and evidence involved in assessing it with respect to nonmaterialist explanations offered by dualists and theists. Dualism and theism are usually discussed in isolation from each other, but both psychologically and logically they go together. I have not conducted a survey, but I suspect that virtually all theists are dualists and almost all dualists are theists, whereas materialists are typically atheists. Theism explains dualism, through God's creation of human souls, so allegedly nonmaterialist aspects of mind such as consciousness and the moral sense provide some evidence for theism. It is natural, therefore, to evaluate the coherence of materialism, theism, and dualism simultaneously. Fortunately, the computational model of coherence developed in chapters 2 and 3 makes this easy to do. What follows is input to the explanatory coherence program ECHO that builds a constraint network and uses the algorithms described in chapter 2 to maximize coherence. This input produces the network shown in figure 4.6. The result of running the program ECHO on this network is that the materialist hypotheses are strongly accepted and the dualist and theistic ones are rejected (figure 4.7). Materialism is more coherent than the combination of dualism and theism.

In the input I have constructed, materialist hypotheses leave unexplained evidence proposition E16, the exis-
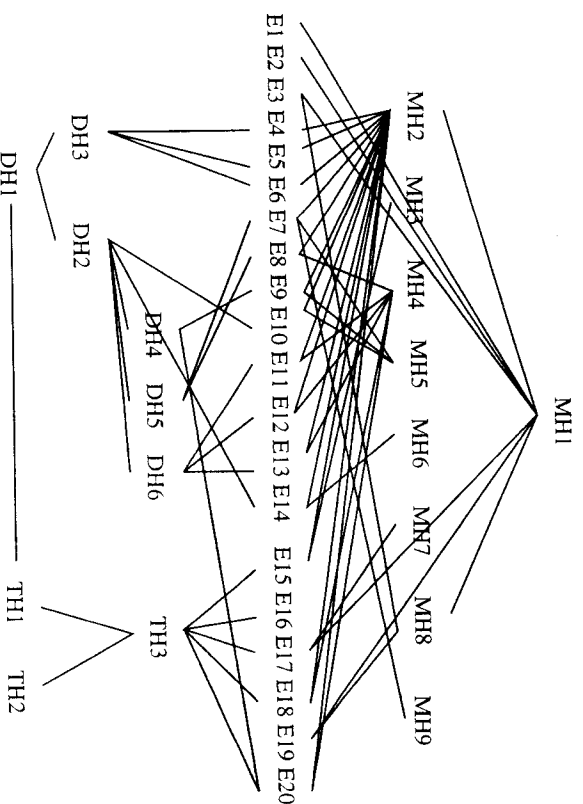
Figure 4.6
The comparative coherence of materialism, dualism, and theism. Lines indicate relations of explanation or implication. Incoherence relations between competing hypotheses are not shown.

tence of universal laws. And theistic hypotheses leave unexplained evidence proposition E19, the suffering of people from diseases and natural disasters. Most important, I have not included a theistic or dualistic explanation of E1 to E3, which are shorthand for the great many physical, chemical, and biological phenomena that science has provided detailed materialistic explanations for over the past several centuries. Overall, the greater explanatory coherence of materialism over the theistic/dualistic alternative is primarily the result of the many theoretical and experimental successes of the sciences.

Many people would disagree with the particular analysis provided in this appendix. My coherence calculation
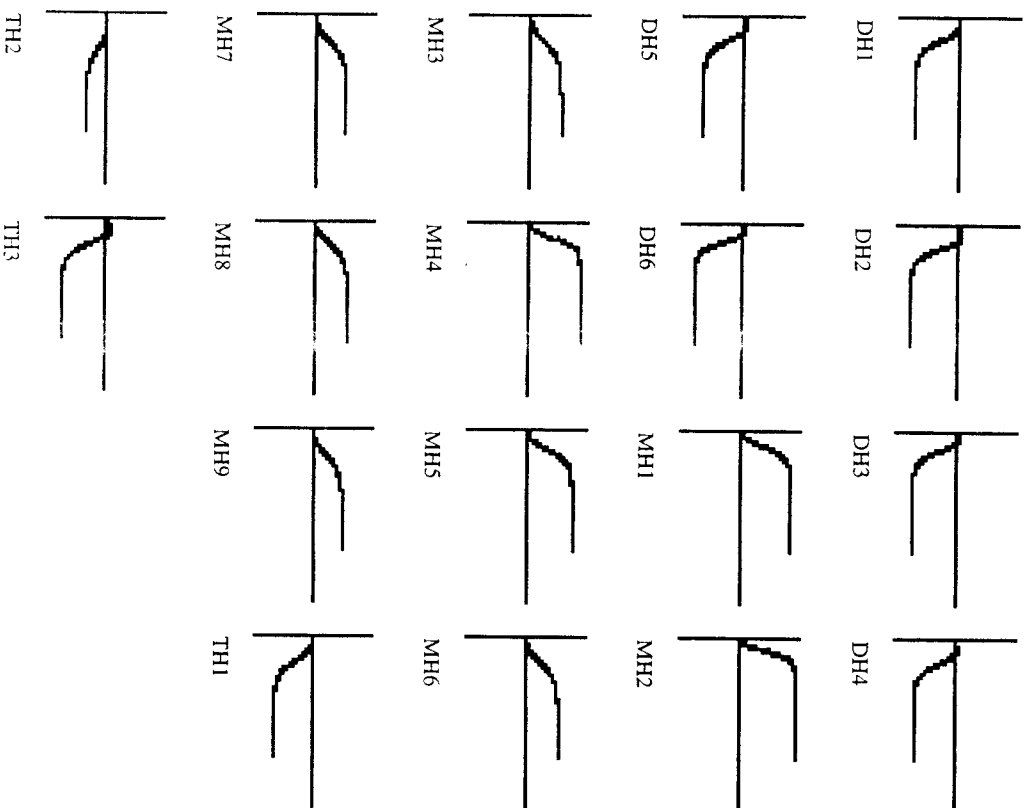
Figure 4.7
Graphs of activation levels of units representing explanatory hypotheses in a connectionist run of ECHO using the input in this appendix. Note that the materialist hypotheses MH1 to MH9 become activated (accepted), while the dualist and theistic hypotheses are rejected.

---

shows only that *if* you accept the input that follows, *then* materialism has greater coherence than its competitors. To dissenters, I recommend the exercise of producing alternative coherence analyses. The main point of this section has not been to provide a definitive refutation of the existence of God, but rather to illustrate how coherence assessments can be applied to metaphysical questions.

Input to ECHO

Materialist hypotheses

(proposition MH1) "Everything consists of matter and energy."

(proposition MH2) "Minds consist of matter and energy."

(proposition MH3) "The universe has always existed, or came to be randomly."

(proposition MH4) "People are prone to fraud, illusion, and other psychological failings."

(proposition MH5) "People acquire beliefs and attitudes through education and socialization."

(proposition MH6) "Consciousness emerges from brain activity."

(proposition MH7) "Biological complexity emerges from natural selection."

(proposition MH8) "People are biological organisms."

(proposition MH9) "Brains near death undergo physical changes."

Dualist hypotheses

(proposition DH1) "Minds consist of matter and soul."

(proposition DH2) "Minds consist partly of soul."

(proposition DH3) "Minds consist partly of matter."

(proposition DH4) "People have free will."

Dualist explanations
(imply (DH1) DH2)
(imply (DH1) DH3)
(explain (DH2) DH4)
(explain (DH2) DH5)
(explain (DH2) DH6)
(explain (DH3) E4)
(explain (DH3) E5)
(explain (DH3) E6)
(explain (DH5) E7)
(explain (DH5) E8)
(explain (DH4) E9)
(explain (DH2) E10)
(explain (DH6) E11)
(explain (DH6) E12)
(explain (DH6) E13)
(explain (DH2) E14)

Theist explanations
(explain (TH1) DH1)
(explain (TH1 TH2) TH3)
(explain (TH3) E15)
(explain (TH3) E16)
(explain (TH3) E17)
(explain (TH3) E18)
(explain (TH3) E19)
(explain (TH3 DH4) E20)

5

Ethics and Politics

In Toronto in 1995, Paul Bernardo was convicted of the prolonged sexual torture and murder of two young women. Since Canadian law does not admit capital punishment, he was sentenced to life in prison. Some people who had long argued the immorality of capital punishment felt strongly inclined to judge that execution would nevertheless be appropriate for Bernardo's extraordinarily heinous crimes. How should such people overcome the incoherence in their ethical views? This chapter shows how justification of ethical principles and particular judgments can be accomplished by taking into account deductive, explanatory, analogical, and deliberative coherence. Like epistemic justification, discussed in chapter 3, ethical justification involves the interaction of several kinds of coherence, with the major addition being the role of deliberative coherence in decision making.

Many ethical theorists have taken coherence to be central to the justification of judgments of right and wrong (Brink 1989, Daniels 1996, De George 1990, DeMarco 1994, Ellis 1992, Hurley 1989, Richardson 1994, Sayre-McCord 1996, Swanton 1992). For example, Rawls writes, "A conception of justice cannot be deduced from self-evident premises or conditions on principles; instead, its justification is a matter of the mutual support of many considerations, of everything fitting together into one

coherent view" (Rawls 1971, 21; see also Rawls 1996, 26, 53, etc.). Unfortunately, ethical theory has remained vague about the nature of coherence and about how ethical principles and judgments can be evaluated with respect to coherence. The term "wide reflective equilibrium" is used to describe a state in which a thinker has achieved a mutually coherent set of ethical principles, particular moral judgments, and background beliefs. But how people do and should reach reflective equilibrium has remained poorly specified. This chapter shows how we can justify ethical principles (such as that capital punishment is wrong) and particular judgments (such as that Paul Bernardo should be executed) by taking into account a wide range of coherence considerations.

To show that ethical decision is a coherence problem, it is necessary to define the elements, the positive constraints, and the negative constraints that operate in ethical thinking. Ethical conclusions require a complex interplay of four different kinds of coherence: deductive, explanatory, deliberative, and analogical. Each of these kinds of coherence involves different kinds of elements and constraints that contribute to an overall conclusion of what ethical principles and judgments to accept. Reflective equilibrium requires integrated assessment of deductive coherence (fit between principles and judgments), explanatory coherence (fit of principles and judgments with empirical hypotheses), deliberative coherence (fit of judgments with goals), and analogical coherence (fit of judgments with other judgments in similar cases).

As I presented it in chapter 1, cognitive naturalism holds that many philosophical issues are intimately connected with the cognitive sciences, including psychology, linguistics, neuroscience, and artificial intelligence. By applying a psychological/computational theory of coher-

ence to ethics, this chapter demonstrates the relevance of cognitive naturalism to ethics.

1 DELIBERATIVE COHERENCE

Standard decision theory says that rationality consists in maximizing the satisfaction of preferences or utilities, but it says nothing about why people have their preferences and utilities. In contrast, Thagard and Millgram (1995; Millgram and Thagard 1996) developed a coherence theory of decision making that involves the evaluation of personal goals as well as actions that potentially accomplish those goals. According to this theory, the elements in deliberative coherence are actions and goals, and the primary positive constraint is facilitation: if an action facilitates a goal, then there is a positive constraint between them. For example, the action of executing Paul Bernardo (or the action of life imprisonment) will facilitate the goal that Paul Bernardo not murder again. Negative constraints arise because some actions are incompatible, since, for example, we cannot both execute Bernardo and imprison him for 50 years. Just as explanatory coherence gives some priority to propositions that state empirical evidence, so deliberative coherence gives some priority to intrinsic goals, ones that an agent has for basic biological or social reasons rather than because they facilitate other higher goals. But just as empirical evidence can be overridden for reasons of explanatory coherence, intrinsic goals can also be revised and overridden for reasons of deliberative coherence, which evaluates intrinsic goals (final ends) as well as instrumental goals and actions. More exactly, deliberative coherence can be specified by the following principles, analogous to those given for explanatory,

deductive, analogical, conceptual, and perceptual coherence in chapter 3:

Principle L1: Symmetry Coherence and incoherence are symmetrical relations: if factor (action or goal) F_1 coheres with factor F_2, then F_2 coheres with F_1.

Principle L2: Facilitation Consider actions $A_1, \ldots, A_n$ that together facilitate the accomplishment of goal G. Then (a) each A_i coheres with G, (b) each A_i coheres with each other A_j, and (c) the greater the number of actions required, the less the coherence among the actions and goals.

Principle L3: Incompatibility (a) If two factors cannot both be performed or achieved, then they are strongly incoherent. (b) If two factors are difficult to perform or achieve together, then they are weakly incoherent.

Principle L4: Goal priority Some goals are desirable for intrinsic or other noncoherence reasons.

Principle L5: Judgment Facilitation and competition relations can depend on coherence with judgments about the acceptability of factual beliefs.

Principle L6: Decision Decisions are made on the basis of an assessment of the overall coherence of a set of actions and goals.

These principles show that deliberative coherence and explanatory coherence have essentially the same structure. Actions are like hypotheses in that they are evaluated with respect to their coherence with each other and with goals that can have a degree of priority on their own, just as evidence can. Both the facilitation relation in deliberative coherence and the explanation relation in explanatory coherence are based on causal connections: actions can cause goals to be satisfied, and hypotheses can state the causes of observations. Despite their isomorphism, however, deliberative and explanatory coherence need to be kept distinct, since the former concerns what to do and the latter concerns what to believe. We could translate a

potential action into a kind of hypothesis, e.g., translate "Execute Bernardo" into the proposition "Executing Bernardo is the best thing to do." But there is no natural translation of goals into evidence, and the facilitation relation that links actions and goals is not the same as explanation, even though both rely on causation: actions do not explain goals. Once networks of elements and constraints are constructed, deliberative and explanatory coherence are computed in the same way, by the algorithms described in chapter 2. But the elements and constraints for deliberative coherence are sufficiently different from those for explanatory coherence that the two kinds of coherence should not be assimilated.

The most novel feature of this account of deliberative coherence is that it allows goals to be evaluated for their coherence with other goals and actions in much the same way as actions are evaluated. Principle L4 assumes that some goals are favored for intrinsic biological or social reasons, but even these goals are evaluated for their overall coherence with other goals. To anticipate an example from chapter 6, hunger may generate the goal of eating from a plate of doughnuts, but other goals such as staying healthy or not looking gluttonous may suppress the goal of eating the doughnuts.

Deliberative coherence is relevant to ethical decisions that take into account the consequences of actions. Someone might argue that executing Bernardo will be cheaper than imprisoning him under special security for life; thus execution facilitates the goal of saving Canadian taxpayers money, unless (as in the United States) the high cost of appeal procedures makes capital punishment more expensive than life imprisonment. The deterrence-based argument for capital punishment also can be reframed as a matter of deliberative coherence: the action of executing murderers facilitates (it is claimed) the goal of preventing

murders. Putting it in this way makes it clear how deliberative coherence depends in part on explanatory coherence. The judgment that an action facilitates a goal depends on a causal judgment about the relation between the action and the goal, and the plausibility of the causal judgment is a matter of explanatory coherence.

In individual decision making, an agent may maximize coherence of actions and goals for the agent alone. Ethical decisions, however, require us to consider what is objectively good, not just for the agent, but also for other people involved. Something is nonmorally good for an agent if and only if it would satisfy an objective interest of the agent (Railton 1986). Normatively, actions should be chosen on the basis of the extent to which they facilitate the objective interests (goals) of all concerned. Thus in deciding whether to execute Paul Bernardo, we take into account the interests of the victims' families, Bernardo himself, and anyone else affected. I am assuming that ethical egoism is false, on the grounds that egoism, or any view that tries to derive ethics only from individual preferences, is incoherent with empirical knowledge about human psychology and sociology and with plausible ethical principles.

Whereas deductive coherence (discussed below) involves a quasi-Kantian concern with general moral principles, deliberative coherence involves a consequentialist concern with goals of those affected by ethical decisions. From the point of view of a coherence theory of ethics, the Kantian and consequentialist positions need not be seen as radically conflicting. Rather, each identifies one kind of coherence that goes into an overall judgment of right and wrong. In everyday debates on ethical issues, people often swing between questions of principle and questions of practical effects. Seeing ethical coherence as involving both deductive and deliberative coherence shows why this can

be so. Deliberative coherence is, however, different from a straightforward consequentialist calculation of the costs and benefits of different actions, because it also assesses the extent to which different goals are important and hence contribute to the assessment of costs and benefits.

Questions of objective interests are closely tied with empirical hypotheses about the wants/interests mechanism of human beings. Evidence from biology, psychology, sociology, and anthropology will be needed to evaluate hypotheses concerning what kinds of actions contribute to the interests of human beings. Thus deliberative coherence is intimately tied with explanatory-coherence evaluation of hypotheses about the nature of humans and their societies. Deliberative coherence does not reduce to explanatory coherence, but depends on it in very useful ways that allow for the possibility of revising views about what is good for people and thereby revising decisions about what to do. For example, the families of Paul Bernardo's victims may naturally want to see him killed, but whether execution would bring some relief from their grief is an empirical question. Without psychological evidence about the effects of executions in similar cases, we do not have grounds for saying whether execution is really in the objective interests of the victims' families.

For utilitarians and other ethical consequentialists, something like deliberative coherence is all there is to ethical decisions. But strict consequentialism generates some implausible judgments, justifying, for example, the horrible mistreatment of a few individuals if it produces the greatest good for the greatest number. Kantian ethics postulates universal principles that establish rights and duties to overrule consideration of consequences of actions. Adoption and application of such ethical principles can be understood in terms of deductive coherence.

2 DEDUCTIVE COHERENCE

For deductive coherence as applied to ethics, the elements are propositions, including both general principles and particular moral judgments. As chapter 3 specified, the main positive constraint is established by the relation of deduction: if one proposition is deducible from another, then there is a positive constraint between them that will tend to make them either accepted together or rejected together. I assume here a psychologically realistic notion of deduction that avoids such trivialities as having a logically contradictory proposition entail every proposition or a logically necessary proposition being entailed by every proposition. In the context of coherence theory, deductive constraints operate quite differently from logical inference, where from $p \rightarrow q$ and p we can infer q by modus ponens, and from $p \rightarrow q$ and not q we can infer not p by modus tollens. As I argued at the end of chapter 3, coherence judgments do not have the kind of step-by-step linear reasoning found in formal logic, but instead require fitting everything together using constraints that are typically soft rather than hard. A soft constraint produces a tendency to accept two positively constrained elements together, but this constraint can be overruled if overall coherence maximization suggests that one of the elements be accepted and the other rejected.

In ethics, positive constraints arise when principles deductively entail judgments, as when the principle that capital punishment is wrong entails that Paul Bernardo should not be executed. Alternatively, the principle that capital punishment is justified for heinous murders implies that Bernardo should be executed. Negative constraints arise because of contradictions between propositions, for example, between the two principles just stated and
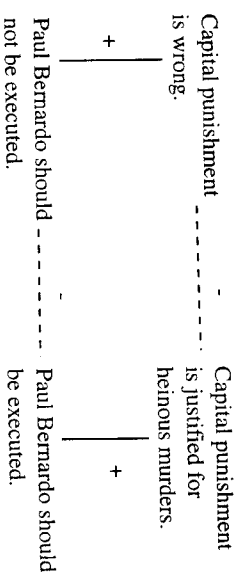

Capital punishment ------- Capital punishment
is wrong.   is justified for
                            heinous murders.

Paul Bernardo should ------- Paul Bernardo should
not be executed.   be executed.

Figure 5.1
Constraint network for the Bernardo case. Solid lines indicate positive constraints, while dashed lines indicate negative constraints.

between the two judgments just stated. Figure 5.1 shows a simple constraint network that shows the relations among these four propositions.

Obviously, the constraint network shown in figure 5.1 does not offer a solution to the coherence problem, since there are two equally coherent solutions: accepting that capital punishment is wrong and that Paul Bernardo should not be executed while rejecting that capital punishment is justified and that he should be executed, or vice versa. Figure 5.1 should be expanded to include higher-level principles such as that killing people is wrong, as well as additional judgments about particular cases of capital punishment. Evaluation of ethical coherence based solely on fit of principles and judgments will generally be open to the standard objection to coherence theories that incompatible sets of propositions can be equally coherent. We will see, however, that broadening ethical coherence to incorporate judgments of explanatory and deliberative coherence can help to overcome this problem by introducing empirical information.

As chapter 3 discussed, deductive coherence is important outside ethics too, for example, in axiom selection in

mathematics. Rarely are axioms selected because they are self-evident. Rather, axioms are selected because they entail the desired theorems, which are in turn accepted because they follow from the axioms. Mathematicians do not proceed from axioms to theorems, nor backwards from desired theorems to axioms; rather they attempt to come up with deductively coherent packages of axioms and theorems. Similarly, ethical principles are not self-evident, but must be selected on the basis of deductive coherence with particular judgments, taking into account additional kinds of coherence.

Particular ethical judgments are also not to be taken as self-evident. Much current ethical theorizing places great weight on intuitions that are established by thought experiments involving hypothetical cases. For example, Thomson (1971) defended the permissibility of abortion by asking you to imagine yourself being kidnapped and having your circulatory system connected to that of a famous violinist in order to save his life. The intuition that you are not obliged to support the violinist for nine months in order to allow his kidneys to recover is then used to support the intuition that abortion is permissible. Cummins (1998) argues convincingly that ethical intuitions as well as other philosophical intuitions derived from thought experiments have little justificatory force, because they are generated from beliefs and tacit theories. Ethical intuitions are thus different from the observations that get a degree of priority in explanatory coherence and from the mathematical intuitions that get a degree of priority in deductive coherence applied to mathematics. Deductive coherence in ethics requires us to find a fit between our general principles and our particular judgments, but I see no reason to give our particular ethical judgments any degree of intuitive priority. Clearly, then, my theory of ethical coherence is not a form of intuitionism. Intuitions

should be viewed not as special inputs to the process of ethical judgments, but as outputs that reflect an overall assessment of what makes sense. Often such outputs have a salient emotional dimension, as chapter 7 discusses.

3 EXPLANATORY COHERENCE

Ethics requires attention to explanatory coherence whenever (as frequently occurs) ethical decisions depend in part on evaluation of empirical hypotheses. Particular judgments such as that Paul Bernardo should be punished depend on factual claims such as that he actually committed the crimes of which he was accused. General principles such as adoption of capital punishment can also be closely tied to factual claims: one common argument for capital punishment is that it is desirable as a deterrent to future crimes, which depends on the empirical hypothesis that having capital punishment as a possible punishment reduces crimes of certain sorts. Evaluation of this hypothesis depends on a very complex evaluation of evidence, such as comparison of countries or states with and without the death penalty. The hypothesis that capital punishment is a deterrent must mesh with a variety of sociological and psychological evidence if it is to be put to ethical use.

How can deductive and explanatory coherence interconnect? The principle that preventing serious crimes is good and the empirical hypothesis that capital punishment helps to prevent crimes together entail that capital punishment is good. These three propositions form a mutually constraining package, as shown in figure 5.2. Unlike a pure deductive principle or moral judgment, however, the empirical hypothesis is subject to a kind of coherence in which empirical evidence is given priority. Priority does not mean that the results of observations must be accepted,
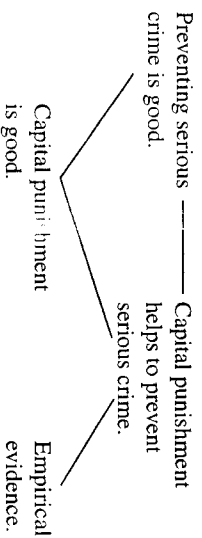
Preventing serious
crime is good. ———— Capital punishment
helps to prevent
serious crime.

Capital punishment
is good.

Empirical
evidence.

Figure 5.2
Deductive coherence depending on an empirical hypothesis. All lines indicate positive constraints based on deductive or explanatory coherence.

only that there is a soft constraint that tends to make them accepted. Now we begin to see how coherence judgments might discriminate objectively between competing sets of principles and judgments: whenever the entailment relation between principles and judgments depends on empirical hypotheses, the coherence of the ethical judgments can be affected by the explanatory coherence of the empirical hypotheses. An opponent of capital punishment might argue that killing innocent people is wrong, and that capital punishment sometimes kills innocent people, so that capital punishment is wrong. This entailment depends on the empirical hypothesis that sometimes innocent people are executed in countries and states that have capital punishment. People who are convinced on the basis of explanatory coherence that the empirical hypothesis that capital punishment sometimes leads to execution of innocent people, and convinced on the basis of explanatory coherence that the hypothesis that capital punishment serves as a deterrent is false, will tend to find more coherent the conclusion that capital punishment is wrong.

From a logical perspective, it might seem odd that if p and q together entail r, then there are pairwise constraints between p and q, between p and r, and between q

and r. However, as in explanatory coherence, these constraints capture the tendency for p, q, and r to fit together as a package of propositions to be accepted or rejected together. Of course, other coherence considerations can lead to some of them being accepted while others are rejected. For a given individual, entailment and explanation relations establish constraints between two elements if the individual believes, on the basis of other coherence judgments, that the relations hold. Logical omniscience and deductive closure have no place in a naturalistic account of inference.

Thus evaluation of ethical principles requires considerations of explanatory coherence as well as deductive coherence and striving for wide rather than narrow reflective equilibrium. But deductive and explanatory coherence are quite similar, in that both involve propositional elements with positive and negative constraints that can be maximized. The interpenetration of deductive and explanatory coherence gives us some hope that ethical deliberation can be affected substantially by empirical evidence. Adding analogical coherence shows another way of broadening ethical coherence.

4 ANALOGICAL COHERENCE

Not all ethical argument considers general principles (deductive coherence) or consequences (deliberative coherence). People often argue for moral principles and judgments analogically, supporting a conclusion in one case by comparing it to a similar case whose moral status is more obvious. The morality of capital punishment is similarly subject to analogical dispute: is execution of a murderer comparable to killing a defenseless victim, or is it somehow similar to acts of self-defense? Applying an analogy to an

Table 5.1
Correspondences between source and target analogs in arguing that capital punishment is wrong

Source	Target
holds(abductor, victim)	holds(state, prisoner)
kills(abductor, victim)	executes(state, prisoner)
wrong(kills(abductor, victim))	wrong?(executes(state, prisoner))

ethical issue requires transferring a moral judgment from an accepted case to a contested case: if capital punishment is relevantly similar to killing a defenseless victim, an act that is obviously wrong, then capital punishment can also be judged to be wrong. Assessing relevant similarity requires establishing correspondences between the source analog, about which an ethical judgment has already been made, and the target analog, to which the ethical judgment is to be applied.

As chapter 3 described, establishing correspondences between source and target analogs can be viewed as a coherence problem involving several different kinds of constraints. The elements are hypotheses about what features of the analogs correspond to each other. Table 5.1 is a simple representation of two analogs. To perform an analogical mapping between these analogs, we need to create mapping hypotheses, such as that *kills* in the source analog corresponds to *executes* in the target analog and that *victim* in the source corresponds to *prisoner* in the target. Once these correspondences are established, analogical inference can support the conclusion that it is wrong to execute a prisoner by mapping *wrong?* to *wrong*.

In the multiconstraint theory of analogy of Holyoak and Thagard (1995), positive constraints are based on

semantic and visual similarity, with people tending to map semantically similar predicates such as *kill* and *execute*. Additional positive constraints are based on syntactic structure: if *holds* in the source is mapped to *holds* in the target, then the corresponding arguments will also be mapped: *abductor* to *state* and *victim* to *prisoner*. Structure also provides negative constraints based on a preference for one-to-one mappings; accepting the mapping hypothesis that *abductor* corresponds to *state* will tend to lead to rejection of the mapping hypothesis that *abductor* corresponds to *prisoner*. Finally, an additional set of positive constraints arises from the purpose of the analogy; what it is designed to accomplish. In ethical deliberations, the purpose of the analogy is to transfer the ethical judgment about the source over to an ethical judgment about the target.

Analogical arguments are rarely convincing on their own, but they can contribute to the overall coherence of a view. Darwin, for example, used an analogy between artificial and natural selection as one of the ingredients in his case for the explanatory coherence of his theory of evolution. Similarly, analogy can help to establish the deductive and deliberative coherence of an ethical conclusion. A defender of capital punishment might argue that just as it may be legitimate to kill an attacker such as Bernardo in self-defense, so it may be legitimate for society to defend itself against murderous psychopaths like Bernardo by executing them. The argument involves both deductive coherence and fit between principles and judgments (killing in self-defense is right; a victim's killing Bernardo would have been justified) and analogical coherence (the comparison between killing for self-defense and execution). Of course, a critic of capital punishment will attempt to undermine this analogy and employ different ones to suggest the applicability of different principles.

5 MAKING SENSE OF ETHICS

From the perspective of the *multicoherence* theory of ethics proposed here, reaching ethical conclusions turns out to be a complex psychological process. Normatively, people can proceed as follows in establishing ethical principles and judgments:

1. Identify deductive elements (principles, judgments) and positive and negative constraints among them.

2. Identify deliberative elements (actions, goals) and positive and negative constraints among them.

3. Identify explanatory elements (hypotheses, evidence) and positive and negative constraints among them.

4. Identify constraints linking the explanatory elements with the deductive and deliberative elements.

5. Identify analogical elements (mapping hypotheses) and positive and negative constraints among them.

6. Identify constraints linking the analogical elements with the deductive, deliberative, and explanatory elements.

7. Finally, use algorithms to maximize coherence by accepting some elements and rejecting others in the way that approximately maximizes satisfaction of the positive and negative constraints.

While this procedure is normatively appealing, it is probably too much to expect of people, given their psychological resources. At the level of consciousness, working memory is far too limited to simultaneously entertain all the different elements that go into such a complex coherence judgment. Perhaps simultaneous maximization of all the constraints goes on automatically at the unconscious level, just as the brain makes sense of complex visual inputs to produce a coherent interpretation of a scene. More likely,

though, the mind must proceed more sporadically, alternating between focusing on one kind of coherence and then focusing on another, or concentrating on some elements and focusing on others (see Hoadley, Ranney, and Schank 1994). Instead of systematically identifying different kinds of constraints, people focus for a while on a particular kind of coherence, such as the deductive fit between principles and judgments, then shift to other kinds of coherence, such as deliberative. Within each focus the mind reaches a tentative coherence conclusion based on the elements and constraints currently active, producing evaluations of elements that can then feed into coherence calculations involving different elements and constraints. This sporadic, unsystematic way of reaching ethical conclusions is obviously subject to the main weakness in any imperfect maximization procedure: instead of reaching a global maximum that achieves the highest possible extent of constraint satisfaction, people may get stuck in a local maximum that, although better than immediately available alternatives, is still inferior to other ways of maximizing constraint satisfaction. One charitable way of explaining the incessant controversies in ethics is by noting the complexity of ethical coherence and conjecturing that disputants have simply fallen into different local maxima.

Those who find inconsistencies in their ethical views, such as the people mentioned at the beginning of this chapter who believe both that capital punishment is wrong and that Paul Bernardo should be executed, can at least attempt to implement the seven-step procedure stated at the beginning of this section. The result should be to bring to bear a wide complex of principles, judgments, actions, goals, hypotheses, evidence, and mapping hypotheses in a way that may suggest how either to abandon the principle that capital punishment is wrong or to reject the judgment that Paul Bernardo should be executed. In either case,

142

coherence with a large number of other considerations will be what determines ethical belief change.

It is important to note that the process by which people reach ethical conclusions is often social: "We press each other toward coherence, and these pressures help nudge us toward consensus" (Gibbard 1990, 204). From the perspective of the individual, it may seem rather arbitrary what elements (concepts, propositions, analogs, etc.) make up the coherence network, but the arbitrariness is much diminished in a social context in which people with different ethical judgments introduce competing elements to be integrated into each other's coherence networks. We do not have to worry about there being an unlimited number of trivial elements that are minor variants of each other, as in Goodman's (1965) "grue" predicates in confirmation theory, so long as ethical coherence is viewed as taking place in human minds in real social contexts. Attention to the content of ongoing controversies should enable us to identify for each ethical issue the relevant elements and constraints. Chapter 7 discusses how consensus can arise through coherence and communication.

6 PUTTING IT ALL TOGETHER

But how do minds amalgamate the various concerns—deductive, explanatory, analogical, and deliberative—that go into an overall coherence judgment? On the traditional view of inference, ethical conclusions would have to somehow integrate the conclusions of a variety of arguments presented one at a time. From the constraint-satisfaction view of coherence, in contrast, inference is not a matter of step-by-step argument, but rather of assembling a set of constraints whose satisfaction is to be maximized in parallel. Figure 5.3 shows how constraints
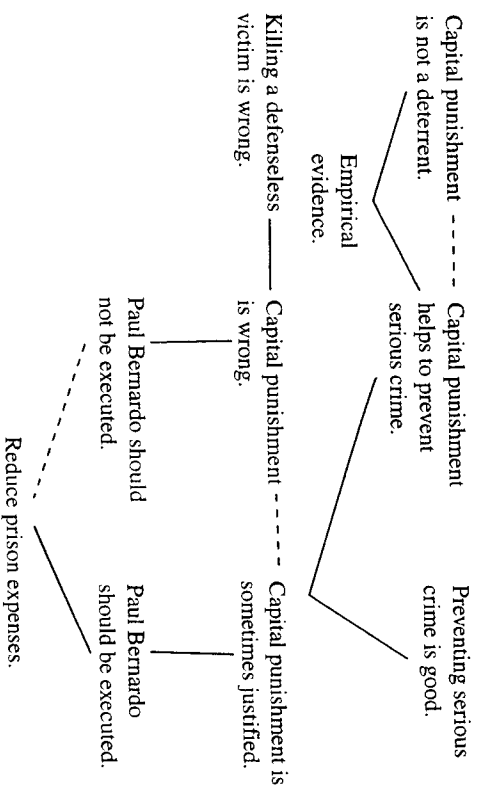
143

Capital punishment - - - - Capital punishment   Preventing serious
is not a deterrent.   helps to prevent   crime is good.
                            serious crime.

          Empirical
          evidence.

Killing a defenseless ——— Capital punishment - - - - Capital punishment is
victim is wrong.   is wrong.   sometimes justified.

              Paul Bernardo should   Paul Bernardo
              not be executed.   should be executed.

                          Reduce prison expenses.

Figure 5.3
Constraint network showing interconnections of explanatory, deductive, analogical, and deliberative coherence. Solid lines are positive constraints, and dashed lines are negative constraints.

derived from deductive, explanatory, analogical, and deliberative coherence can all be incorporated into a single constraint network. The figure shows only a fraction of the considerations that would go into a full assessment of the morality of capital punishment, but it serves to show how a wide variety of constraints can be incorporated into a single network. I have run a computer simulation using the following programs together to produce a common network:

• ECHO (Thagard 1992b) creates constraints based on whether the hypothesis that capital punishment is a deterrent explains the evidence. ECHO is also used to approximate the deductive relation that justifies capital punishment as following from the principle that preventing serious crime is good, as well as other deductive

relations such as the one between "Capital punishment is wrong" and "Bernardo should not be executed."

• ACME (Holyoak and Thagard 1989) creates constraints based on the analogical mapping between capital punishment and killing defenseless victims.

• DECO (Thagard and Millgram 1995) creates constraints based on consequences such as that capital punishment reduces prison expenses.

Because ECHO, ACME, and DECO all use the same connectionist algorithm for maximizing coherence (chapter 2), the computer simulation succeeds in reaching a conclusion based simultaneously on all the considerations shown in figure 5.3.

Of course, this simulation does not settle the enormously difficult ethical issue of whether capital punishment is justified. It does serve, however, to show how various kinds of coherence considerations can combine to produce an overall judgment. To fully capture individual judgments about capital punishment, it would be necessary to combine assessment of the ethical issue with metaphysical views of the sort discussed in chapter 4. People who believe that God ordains that murderers be put to death will obviously reach a different conclusion than others whose coherence calculations are restricted to secular matters.

7 THE COHERENCE OF ABORTION

The complexity of ethical coherence is further illustrated by debates concerning the morality of abortion, which contain a variety of deductive, explanatory, deliberative, and analogical considerations. Baird and Rosenbaum (1993) contains the U.S. Supreme court judgment on *Roe*

v. Wade, which is clearly based on a mixture of coherence considerations, along with various essays for and against abortion that also illustrate the multifariousness of coherence. Deductive arguments are used by both sides of the issue. Defenders of abortion argue that the illegitimacy of the state's banning abortion follows from a right to privacy, whereas critics of abortion claim that its immorality follows deductively from the principle that murder is wrong. Of course, neither of these deductive arguments is convincing to the other side, since they depend on the legitimacy of the principle stated and on the acceptability of additional premises required to make the argument sound, for example, that abortion is murder.

Other arguments invoked in the abortion case point to issues of deliberative coherence, for example, on the pro side that prohibiting abortion will lead to injuries of numerous women undergoing illegal abortions, and on the con side that abortion causes distress both to fetuses and to women who have abortions. As in the issue of capital punishment, deliberative coherence often interacts with explanatory coherence: factual claims such as that making abortion illegal would cause suffering both from illegal abortions and unwanted children need to be empirically evaluated on the basis of how well they fit with theories and observations in psychology and sociology. Explanatory coherence can also interact with deductive coherence, for example, when theists infer that abortion is wrong because God forbids it. This deductive argument presupposes that there is a God, a hypothesis that can be evaluated on the basis of its explanatory coherence.

Analogical coherence also enters into judgments about the morality of abortion, since arguments often rely on comparison to practices such as infanticide or hypothetical cases such as described earlier of the one being involuntarily connected with a violinist. Analogy also plays

a major role in legal judgments, when abortion is treated as a case that should be settled in ways similar to precedents such as the judgment that prevented states from banning contraception. Whether abortion is deemed as analogous to legally acceptable practices such as contraception or as analogous to proscribed practices such as infanticide depends on a variety of deductive, explanatory, and deliberative considerations. Analogies contribute to the assessment of explanations and actions, just as the assessment of explanations and actions contributes to the evaluation of analogies. There is no circularity here, because all kinds of coherence can be simultaneously computed by global maximization of satisfaction of different kinds of constraints.

Thus, like capital punishment, the ethical assessment of abortion depends on a combination of deliberative, explanatory, deductive, and analogical coherence. I have attempted not to provide such an assessment here but only to indicate how the different kinds of coherence combine to influence judgments about the morality of abortion.

8 NORMATIVE ISSUES

My theory of ethical coherence is intended to be both descriptive and prescriptive, characterizing how people think ethically when they are thinking at their best. Epistemology can be "descriptive," i.e., simultaneously descriptive and prescriptive (Thagard 1992b, 97). But linking the descriptive and the normative is more problematic in ethics than it is in epistemology. In philosophy of science, we can take as exemplars of scientific inference those scientists who have made the most important contributions to the growth of knowledge: Newton, Darwin, Einstein, and so

on. In ethics we do not have recognized inferential experts whom we can view as exemplars. The most influential ethical theorists have tended to be dogmatic in defending monolithic approaches to ethics, for example, exclusively in terms of Kantian rights and duties or exclusively in terms of utilitarian consequences. My more eclectic coherence approach makes possible incorporation of a wide variety of ethical considerations but must face the question of whether putting them all into a coherent soup will produce judgments that are objectively right. My response is that when dealing with difficult ethical issues such as capital punishment and abortion, we should feel obliged to take into account all the different kinds of issues that have been taken to be relevant to the morality of such practices. One of the advantages of the coherence theory is that it can incorporate the full range of arguments that ethicists have used in a more piecemeal fashion to support their own conclusions. The requirement of taking into account a broad range of considerations is analogous to the requirement in epistemology that anyone evaluating an hypothesis on the basis of its explanatory power should take into account the full range of empirical evidence and alternative hypotheses.

Still, difficult normative issues arise in the application of ethical coherence. We saw in the discussion of empirical issues in ethics that explanatory coherence can affect deliberative coherence when judgments of likely consequences of actions are based on causal theories and evidence. In a coherence system, however, there is a danger that deliberative coherence will have an undesirable effect on explanatory coherence, as when people adopt hypotheses for personal gain rather than on the basis of evidence. There is substantial psychological evidence that people's goals do affect their evaluation of evidence (see Kunda 1990).

Normatively, however, we want explanatory coherence to affect deliberative coherence and not vice versa. This issue is addressed further at the end of chapter 6.

Another difficult normative issue concerns the construction of constraint networks, such as the one shown in figure 5.3. Different people may put different weights on the positive and negative constraints connecting the various elements in the network. At the extremes, a devout Kantian might put zero weight on any empirical considerations, and a utilitarian might put zero weight on anything else. My first response is to point out that the extreme versions of both these approaches have familiar incompatibilities with most people's ethical judgments: Kantian rules such as never to tell lies are too rigid to apply universally, and utilitarian calculations that count the pleasure and pain of strangers equally with the pleasure and pain of loved ones are impossible for most people. My second response is to point to the multifarious nature of actual ethical arguments that embrace different kinds of ethical concerns, including both Kantian and utilitarian ones. I do not have an algorithm for establishing the weights on people's constraints, only the hope that once discussion establishes a common set of constraints, coherence algorithms can yield consensus. I return to the topic of consensus in chapter 7. Now I turn to the discussion of important normative issues in politics.

9 POLITICS: JUSTIFYING THE STATE

The term *politics* is usually defined as the art or science of government, so the first normative political issue is whether there should be any government at all. Do organized nation states have legitimate authority over their members, or is government an illicit infringement on the

freedom of people forced to submit to the decrees of a state? Anarchists, who advocate the abolition of government, claim both that there is no justification for the state and that its elimination will produce a society in which people's lives are improved. Traditional anarchists like Mikhail Bakunin and Peter Kropotkin advocated elimination of the state in favor of a cooperative socialism in which people would provide mutual aid. In contrast to this sort of left-wing anarchism, there is a more recent and currently more popular brand of right-wing anarchism, which advocates full-fledged free-market capitalism as the alternative to current government (Sanders and Narveson 1996). At the extremes, left- and right-wing thinkers converge on rejection of the state, although the utopian forms of government-free life that they envision are very different.

What response can one give to the rejection of the state by various anarchists? Foundationalists, who think that politics, like epistemology and ethics, requires indubitable truths, must find incontrovertible axioms and implications that lead to the conclusion that the state is justified. But such a foundation is no more likely to be found in politics than in epistemology or ethics. Instead, we need to look for a coherentist justification of the state that combines deliberative, analogical, explanatory, and deductive considerations.

In particular, the question of whether people should live in a state or in socialist or capitalist anarchy is largely a matter of deliberative coherence. At a crude level, here are the actions to choose from:

• Establish a nation state that has authority over its citizens.

• Abolish the state in favor of socialist cooperation.

• Abolish the state in favor of capitalist free markets.

Both left- and right-wing anarchists assume that people will be better off without the state, but in what respects? What are the goals with respect to which the deliberation concerning the existence of the state should take place? It is impossible to establish deductively the goals that political deliberation should accomplish, but three stand out prominently in political arguments. I shall call them the F-*constraints*: freedom, flourishing, and fairness. Deliberation both about whether there should be a state and about what form the state should take can be framed in terms of how different ways of organizing people contribute to satisfaction of these three constraints. Other kinds of coherence, particularly analogical and explanatory, will interact with deliberative coherence to produce a coherentist justification of particular forms of the state.

Without intending to give it any kind of priority, I listed freedom as the first F-constraint. Freedom (autonomy, liberty) is the ability of individuals to make personal and economic decisions without interference by the state or other people. Initially, it might seem that anarchism is clearly the way to maximize freedom, but eliminating the state may in fact increase interference by other people, who are also unconstrained by the state. We need to weight carefully the relative extent to which the options of (1) having a state, (2) right-wing anarchism, and (3) left-wing anarchism promote freedom.

Similarly, we need to weight the contributions that different forms of government and nongovernment make to human flourishing, which encompasses both happiness and excellence. People flourish not only when they enjoy pleasure and lack pain, but also when they accomplish things that humans do best, including intellectual accomplishments (such as science, philosophy, and art) and physical accomplishments (such as athletics). Anarchists from both the right and left assume that people will not only

have more freedom without the state, they will also flourish more without state interference. Evaluating these assumptions requires explanatory and analogical considerations, described below.

The final F-constraint is fairness, which concerns the extent to which there is equality in the distribution of freedom and flourishing. Consider a society in which most people enjoy great degrees of flourishing and freedom at the expense of some people who are totally deprived of these benefits, perhaps because they are slaves to the well-off. Such a society is so unfair that most people would consider that it is not justified by the freedom and flourishing it provides for those who are well-off. I see no way in which the F-constraints can be subordinated to each other, even though they have interactions. Freedom, for example, seems historically to contribute to flourishing, but that does not mean that is valuable only for its role in promoting flourishing. In different historical contexts, there may be varying ways in which the different F-constraints enhance or weaken each other's satisfaction.

Different traditions in political and social philosophy have emphasized different constraints. For libertarians, the primary constraint is freedom from interference by others, with little concern for flourishing or fairness. For utilitarians, the only constraint on ethical and political justification is maximizing the greatest happiness for the greatest number, which falls under my constraint of flourishing. Various theorists, from socialists and anarchists to liberals such as John Rawls, have stressed fairness as a key constraint on any admissible political system. My view is that we should take all three of these constraints seriously in trying to justify the state and particular versions of it. It is an open and difficult question what the relative weights of these constraints should be; I conjecture that differences in political philosophy arise primarily from different

weightings of the importance of freedom, flourishing, and fairness.

How can we use freedom, flourishing, and fairness to compare having a state versus left- and right-wing anarchism? The answer involves both analogical and explanatory coherence. Explanatory coherence is relevant to assessing the causal claims that underlie facilitation relations that produce judgments of deliberative coherence. We need to assess the claim that eliminating the state would facilitate freedom, flourishing, and fairness. Unfortunately, there is no evidence that we can use to assess the explanatory coherence of this claim, because complex human societies have lived under some form of government at least since around 2800 B.C., when Sumerian city states were in operation. There simply is no evidence that life without the state has facilitated or would facilitate freedom, flourishing, and fairness.

Anarchism also fares poorly when analogical coherence is taken into account. Even though there have been no state-free episodes to establish facilitation relations directly, one might argue analogically that a future stateless society might have the good features of some past situation that was less state-dominated than current societies. But what are the analogs one can use? For right-wing anarchists, perhaps the best choice would be capitalist governments before the twentieth-century rise of the welfare state increased state involvement. But it would be very hard to make the case that nineteenth-century residents of Britain or the United States had lives superior to people in those countries today. Perhaps there was more abstract economic freedom than now exists, but health and education—crucial ingredients in flourishing—were far inferior to current standards. Moreover, fairness was intensely violated by the great discrepancies in political participation (the right to vote was limited) and wealth. The claims of

right-wing anarchists to satisfaction of the F-constraints are thus bereft of explanatory and analogical coherence. Of course, they may well claim that the only constraint that matters to them is freedom, so that failures of right-wing anarchism to afford gains in flourishing and fairness are irrelevant. But the restriction to only one constraint is arbitrary and insupportable. Anarchism may follow deductively from some principle that says that freedom is all that matters, but that principle does not cohere with what we know about human needs and desires. Myopic concentration on freedom is no more appealing than the opposite view that freedom may be largely ignored in order to increase general flourishing and fairness.

Left-wing anarchists are even shorter on plausible analogs than right-wing anarchists. For examples of stateless societies run on principles of cooperation and social aid, one can look only to relatively small groups, such as communes and Israeli kibbutzim. However, there are two reasons why these analogs do little to support the claim that socialist anarchism would support freedom, flourishing, and fairness. First, the analogy between anarchist experiments in small groups and running a complex society without a state is very weak: there are huge differences between running a society with millions of people and running a group of twenty or one hundred people. Second, even on a small scale, anarchistic, socialistic experiments have not been very successful in the long run. The Israeli kibbutz movement was strong in the 1950s, but today there are only weak remnants trying to survive as quasi-capitalist enterprises. Other anarchist experiments have degenerated into chaos or despotism. So analogies do little to support left-wing anarchism.

Therefore, in order to facilitate freedom, flourishing, and fairness, having some form of government is preferable to having no state at all. Of course, states have varied

154

enormously in the degree to which they satisfy these constraints, which raises the question of what kind of government is best. Given the evidence that the best of modern states contribute substantially to freedom, flourishing, and fairness and the lack of evidence that anarchism in any form would even come close to performing so well, we can dispense with the skeptical question of whether the state is justified at all and move on to the much more interesting and important question of what kind of state is best.

10 WHAT KIND OF STATE?

The question of whether the state is justified is of purely philosophical interest, since hardly anyone seriously considers the complete abolition of the state. But the question of what kind of state to have is very much alive in many contexts. For example, in the wake of the collapse of communism, people in Eastern European countries are faced with deciding what kind of government should replace it. Should they move towards a kind of laissez-faire capitalism at the opposite extreme from socialism, or should they look for a middle road closer to social democracies such as Sweden, or should they revert to a version of socialism without the extreme restrictions on freedom found under communism? Choices in most Western states are less extreme, ranging between social democracy and welfare capitalism in Western European states and Canada, and between welfare capitalism and laissez-faire capitalism in the United States. We can now explore how freedom, flourishing, and fairness fare in various societies to help answer the question of what kind of state is best.

Deciding what kind of state to adopt is primarily a matter of deliberative coherence subject to the three F-constraints. But what are the options? Derbyshire and

155

Derbyshire (1997) provide a systematic comparison of 192 current states, which they classify into the following political systems:

• Liberal democracy, with representative government and individual freedom, e.g., the United States
• Emergent democracy, like liberal democracy, but with limited political stability, e.g., Poland
• Communism, with state ownership, and one-party control, e.g., China
• Nationalistic socialism, with charismatic leaders, e.g., Libya
• Authoritarian nationalism, with one-party dominance, but not socialist, e.g., Indonesia
• Military authoritarianism, e.g., Nigeria
• Islamic nationalism, e.g., Iran
• Absolutism, with no constitutional government, e.g., Saudi Arabia

If these are the eight options for choosing a kind of state, then choice is relatively easy. The 73 liberal democracies not only surpass the other states in freedom, they also by and large have much greater degrees of flourishing, as measured by such variables as wealth, health, and education. As for fairness, the liberal democracies make voting generally available, and their distribution of wealth is generally no worse than that of other kinds of state. Hence liberal democracy is clearly superior to all other current forms of government with respect to the F-constraints.

Choice gets more difficult if we try to select among different variants of liberal democracy. We can distinguish at least the following variants, distinguished by the increasing extent to which the state is involved in the economy:

- Laissez-faire capitalism, e.g., nineteenth-century Britain
- Welfare capitalism, e.g., Britain since the Second World War and the United States since Roosevelt's New Deal
- Social democracy, e.g., Sweden

To decide which of these to prefer, we need to make a much more fine-grained assessment of freedom, flourishing, and fairness. Ideally, we would need to conduct a full survey with informative measures of the degree to which current countries and ones in the recent past have satisfied the F-constraints. No such surveys currently exist, but there have been other surveys that address some of the relevant issues.

For a start, the Fraser Institute, a Canadian economic think tank, publishes an index of economic freedom, which attempts to measure one aspect of freedom. This index aims to measure the extent to which individuals are free to choose for themselves and engage in voluntary transactions with others, and have their rightly acquired property protected from invasions by others (Gwartney and Lawson 1997, 2). The index contains seventeen components, divided into four major areas:

- Money and inflation: protection of money as a store of value and medium of exchange
- Government operations and regulations: freedom to decide what is produced and consumed
- Takings and discriminatory taxation: freedom to keep what you earn
- Restraints on international exchange: freedom of exchange with foreigners

The political bias of this way of measuring freedom is evident: it is part of the Fraser Institute's mission to reduce taxation and other forms of government involvement in

Table 5.2
Summary rankings of the economic-freedom ratings by the Fraser Institute, 1997, showing the top twenty countries

Rank	Country	Freedom rating
1	Hong Kong	9.6
2	Singapore	9.4
3	New Zealand	9.2
4	United States	9.1
5	United Kingdom	9.0
6	Canada	8.8
7	Argentina	8.7
8	Netherlands	8.6
8	Panama	8.6
8	Australia	8.6
8	Luxembourg	8.6
8	Ireland	8.6
13	Switzerland	8.5
14	Japan	8.3
14	Denmark	8.3
14	Norway	8.3
17	Belgium	8.2
17	El Salvador	8.2
17	Finland	8.2
17	Germany	8.2

Source: Gwartney and Lawson 1998, p. 22.

the economy. Although this measurement of economic freedom is not an adequate substitute for the freedom constraint, it is a methodologically interesting way of beginning to quantify ideas about freedom. Table 5.2 reproduces part of the Fraser Institute's 1998 ratings of economic freedom.

Aspects of flourishing can also be measured with some degree of approximation. Although it is in many ways limited as an indicator of human flourishing, the United

Nations Human Development Index (HDI) provides an interesting first approximation. The HDI is a composite of three basic components of human development:

- Longevity, measured by life expectancy
- Knowledge, measure by a combination of adult literacy and mean years of schooling
- Standard of living, measured by purchasing power, based on real GDP per capita adjusted for the local cost of living

Table 5.3 lists the top finishers in the most recent assessment. There does not appear to be any strong correlation with the economic-freedom-index results shown in table 5.2. Nor can we demonstrate that the three components of the HDI correlate strongly with human happiness and achievement of excellence, although it is not implausible that they do. But the methodology of the HDI shows that it is in principle possible to assess the extent to which different countries have enabled their citizens to flourish. By extension, once the countries are classified according to what kind of government they have, we can begin to assess the contributions of different kinds of states to human flourishing. In contrast to the economic-freedom tally, which was dominated by states inclined toward laissez-faire policies, states with relatively more state intervention in the economy and social planning tended to do well according to the human development index.

One major weakness in both the HDI and the economic-freedom index is that they look only at aggregates and neglect questions concerning the distribution of freedom and flourishing. The United Nations does, however, offer measures of poverty and gender inequality that address these issues to some extent. A fairness index needs to be developed to measure the extent to which there is an equitable distribution of economic and social goods not limited by gender, race, and ethnicity. If the aim of this

Table 5.3
Top twenty countries in the 1997 United Nations Human Development Index

HDI rank	Country	Life expectancy at birth (yrs.) 1994	Adult literacy rate (%) 1994	Combined 1st, 2nd, 3rd level gross enrollment ratio (%) 1994	Real GDP per capita (PPP$) 1994	Human Development Index (HDI) 1994
1	Canada	79.0	99.0	100	21,459	0.960
2	France	78.7	99.0	89	20,510	0.946
3	Norway	77.5	99.0	92	21,346	0.943
4	USA	76.2	99.0	96	26,397	0.942
5	Iceland	79.1	99.0	83	20,556	0.942
6	Netherlands	77.3	99.0	91	19,238	0.940
7	Japan	79.8	99.0	78	21,581	0.940
8	Finland	76.3	99.0	97	17,417	0.940
9	New Zealand	76.4	99.0	94	16,851	0.937
10	Sweden	78.3	99.0	82	18,540	0.936
11	Spain	77.6	97.1	90	14,324	0.934
12	Austria	76.6	99.0	87	20,667	0.932
13	Belgium	76.8	99.0	86	20,985	0.932
14	Australia	78.1	99.0	79	19,285	0.931
15	U.K.	76.7	99.0	86	18,620	0.931
16	Switzerland	78.1	99.0	76	24,967	0.930
17	Ireland	76.3	99.0	88	16,061	0.929
18	Denmark	75.2	99.0	89	21,341	0.927
19	Germany	76.3	99.0	81	19,675	0.924
20	Greece	77.8	96.7	82	11,265	0.923

Source: http://www.undp.org/hdro/. "PPP" stands for "purchasing power parity."

section were to argue for a particular form of state, I would need to attempt to quantify the extent to which different countries satisfy the fairness constraint. My aim, however, is more methodological: to show that in principle we can assess different kinds of states with respect to the extent to which they satisfy the F-constraints. Although the assessment is obviously a very challenging project in social science, and although the tough issue of how to weight the

constraints of freedom, flourishing, and fairness remains unsolved, we can at least begin to see how the problem of justifying particular forms of states can be seen as a coherence problem.

It is unusual for people to undergo major changes in their political views during their lifetimes, but it sometimes happens. Consider, for example, the student radicals of the 1960s who abandoned the traditional liberal democratic views that they grew up with in favor of more revolutionary ones. Or consider the neoconservative intellectuals of the 1970s and 1980s, some of whom had been much more radical in their youth. What is involved in the shift from being a liberal to a left-wing radical, or from a leftist to a neoconservative espousing the virtues of laissez-faire capitalism? Evidence to answer this question is limited, but the coherence perspective suggests that we should look at changes such as the following:

- Changes in beliefs about human nature, based on explanatory and analogical coherence

- Changes in beliefs about the efficacy of different political strategies, again based on explanatory and analogical coherence

- Changes in the weights attached to the F-constraints, altering the relative priority given to freedom, flourishing, and fairness

It is difficult to say to what extent the latter process is a rational one. Emotional changes of the sort discussed in chapter 6 will also be relevant.

Choosing what kind of state to adopt is largely a matter of deliberative coherence with the F-constraints, but analogical reasoning can also contribute. Negative analogies are particular states that we do *not* want future states to be like, for example, Nazi Germany and the Soviet

Union under Stalin. Positive analogies are particular states that have aspects that we might want to emulate, for example, the freedom of the United States and the fairness of the Scandinavian social democracies. Analogies may myopically limit deliberative coherence, since they focus on past examples rather than novel future state organizations that surpass previously available ones, but they can provide positive and negative suggestions about what to keep and what to avoid in designing the state. As in the assessment of the ethical coherence of capital punishment, explanatory coherence becomes relevant to assessing the plausibility of relevant empirical claims concerning the efficacy of different kinds of political organization. For example, the claim that a particular kind of state promotes prosperity must be evaluated against the historical evidence. Hence, justifying the state and the more specific task of choosing what kind of state to adopt are both coherence problems.

11 CONCLUSION

This chapter has proposed a multicoherence theory of ethical thinking according to which people reach ethical and political conclusions by approximately maximizing the satisfaction of deductive, explanatory, deliberative, and analogical constraints. There are at least four reasons why this theory should be adopted as a normative account of how people should reason about right and wrong.

First, the multicoherence theory of ethics can handle the complexity of moral reasoning. This chapter has shown the relevance of all four kinds of coherence to the evaluation of whether capital punishment and abortion are right or wrong. It would not be hard to show that other major ethical issues similarly involve a mixture

of deductive, explanatory, deliberative, and analogical considerations.

Second, the multicoherence theory is naturalistic in that it is consistent with substantial amounts of evidence showing that the processes of parallel constraint satisfaction are important in human cognition (Holyoak and Spellman 1993; Thagard 1996, chap. 7). If vision, language understanding, hypothesis evaluation, concept application, and analogy are all coherence processes, it should not be surprising that ethical thinking is also a coherence process. The ethical theory developed in this chapter is not naturalistic, however, in the sense of claiming that moral judgments are reducible to scientific facts about the natural world. Judgments about right and wrong are often closely tied with scientific judgments, as we saw with the interconnections among deductive, deliberative, and explanatory coherence. But these connections do not *reduce* deductive and deliberative coherence to explanatory coherence. Ethical questions are not simply factual questions, but they are sufficiently linked with empirical issues that we can hope that agreement on psychological, biological, and economic issues can contribute to agreement on ethical issues. My multicoherence account of coherence provides a much fuller account of ethical inference than is found in recent naturalistic accounts that emphasize either perceptionlike neural networks (Churchland 1995, Flanagan 1996) or metaphor (Johnson 1993, 1996; Lakoff 1996). These accounts capture aspects of conceptual and analogical coherence, but neglect the contributions of deductive and deliberative coherence to ethical judgments.

Third, the coherence view of ethics and politics proposed here avoids the two major problems of foundationalist approaches to ethics and epistemology. The first problem is that, for epistemology as for ethics, no one has ever been able to find a set of foundations that even comes close to

receiving general assent. The coherentist approach has no need for a priori intuition or contractarian artifice. The second problem is that proposed foundations are rarely substantial enough to support an attractive epistemic or ethical edifice, so that foundationalism degenerates into skepticism. In contrast, the multicoherence theory of ethics, like coherence theories of knowledge, recommends that we jump into issues in midstream, revising ethical beliefs as necessary to increase overall coherence, without attempting the impossible task of rederiving all ethical principles and judgments from first principles.

Finally, the multicoherence theory proposed here has the advantage over previous coherentist approaches to ethics that it employs a clearly stated and computationally implemented account of what it is to maximize coherence. Explanatory, deliberative, and analogical coherence all have computational models that have been applied to numerous complex real-world cases. Amalgating these kinds of coherence with the deductive coherence of ethical principles and judgments is nontrivial, but the sporadic, incremental way in which people generally shift focus among different kinds of coherence can be seen as a rough approximation to a more ideal process of global maximization of constraint satisfaction. We do not always maximize coherence, but sometimes we manage nevertheless to make quite good sense of right and wrong.

I have so far neglected an important psychological aspect of ethical thinking. When people make ethical and political judgments, there is usually a strong emotional component. People feel very positively about what they view as right, and they feel strong negative emotions about what they view as wrong. Chapter 6 develops a theory of emotional coherence that shows how to integrate the coherence considerations discussed in this chapter with emotional matters.

12 SUMMARY

In contrast to the vague notions of coherence used by many ethical theorists, the theory of coherence as constraint satisfaction can provide a detailed and computable model of how different kinds of coherence can contribute to ethical judgments. Deliberative coherence involves choosing actions and goals on the basis of their coherence and incoherence with other actions and goals. Deliberative coherence is essential to ethical judgments, but deductive, explanatory, and analogical coherence can also contribute. This theory of ethical coherence is intended to be both descriptive of how people make ethical judgments and prescriptive of how they should. Political judgments involving the justification of the state and the choice of a kind of state are based on ethical coherence, particularly on deliberative coherence with respect to the goals of freedom, flourishing, and fairness.

6

Emotion

Like most philosophical and psychological writings about inference, my discussion of coherence has so far ignored the important role of emotion in human cognition. This chapter presents a theory of emotional coherence and describes its implementation in a computational model that has been applied to interpersonal trust and other important psychological phenomena that involve both inference and emotion, including empathy and nationalism. The theory and model are then extended to encompass "metacoherence" and the emotional impact of overall assessments of coherence relevant to understanding beauty, humor, and cognitive therapy.

1 THE IMPORTANCE OF TRUST

When Jimmy Carter ran for President in 1976 in the wake of Watergate, he told the voters, "You can trust me." After Tony Blair was elected Prime Minister of England in 1997, he responded by telling the voters, "You have put your trust in me. I will not let you down." In elections, politicians often try to convince the voters that they are more trustworthy than their opponents, and incumbents work to maintain the trust of their constituents (Bianco 1994, Fenno 1978). The political importance of trust is