Fact-Finding
An Introduction to Critical and Scientific Thinking

Eduardo Wilner

Department of Philosophy & Humanities

Department of Biology and Wildlife

University of Alaska, Fairbanks

P. O. Box 755740

Fairbanks, AK

99775-5740

(907) 474-7343

emwilner@alaska.edu

Table of Contents

THE GOALS

One of the main mandates of the Alaska BioPREP program is to enhance the early-college level of understanding and appreciation of molecular biology in particular, and (necessarily) of science in general. One of the main strategies for this is the production of hands-on projects involving state-of-the-art experimental settings for students to explore as they learn.

This is a good strategy. However, as most science educators know, even those students that are engaged by these experiences have a hard time taking the knowledge they acquired in the science classroom and making it bear on their general grasping of the world. Or, to put it slightly differently, they have a hard time taking science seriously: appreciating scientific understanding to the point in which they feel compelled to bring it to bear on their understanding of the world in general.

It is our experience that students are able to do this, to devalue science to the point that it's easily tucked away when they are outside of the science classroom, thanks to their deeply-held but mistaken assumption that knowing through science, and knowing through experience in general, are two completely different types of things. Thus, failing to understand how the scientific way of knowing in particular relates to ways of knowing in general allows students to devalue science. One of the main objectives of this guide is to tackle this misconception and thus make science relevant.

This guide is written in a conversational tone and with a reader in mind that, albeit eager to learn and think with the author, is not necessarily well versed in the nature of critical thinking in particular, or of science in general. Thus, I have strived to provide sources and extra-reading material that are accessible to the lay reader and, when possible, available for free in the Internet.

Now relax, read, and enjoy.

1. CHAPTER 1 -- FACT-FINDING

1.1 A General Introduction

This book is about facts. It is not so in the sense that you're about to read a long list of facts out there. This book is about what facts are, and how can we find them. In other words, we must start with considering the question of what we might, and could possibly mean when we say that x, y, or z are *facts.*

Questions such as these are dangerous. Ideas that lay at the very foundations of our thinking are fearful things to mess with. Think about how exited would you be if I propose we start messing with your house's foundations. In this very architectural sense, we are going to call those ideas that support all the other ideas in the scaffolding of our minds (rather uninventively) "foundational." A good way to test if an idea is foundational or not, would be to shake it, sort to speak, and to see if most of the rest of our conception of the universe also shakes. To illustrate this, let's get science-fictiony for a second.

Imagine that some evil genius puts your brain in a vat, connected not only to a capable life support system, but also to a futuristic virtual reality computer. This virtual reality generator is so good, that you cannot possibly tell the difference between its virtual production, and the real thing. Thus trapped in her universe, she can mess with any of your ideas, manipulating them and changing them at will. If she were to do the most damage to your mental universe, she would aim (you guessed it) to its foundations. For instance, she would fuddle the distinction between *fact* and *whimsical opinion*. Is arsenic good for you? Does gravity hold? Am I really here? Am I me? The deeper we go into the foundations, the more terrifying it gets.

I'm sorry to say that, even though I'm no genius (evil or otherwise) the purpose of this book is to investigate a dreadful subset of all possible foundational questions: the family of ideas that tackle the question of human knowing. For instance, consider a really obvious, unquestionable fact. Let's say: "You are reading this book now." Questioning this sounds beyond silly, doesn't it? But let's do it anyway. First of all: what do we mean when we call this a fact? A cloud of words comes to the rescue. As we already saw, we mean that it's *unquestionable*. We mean that it is *true* - not "true for me," but true with a capital "T." A fact, the real deal, is a fact for everybody. Fine, let's dig deeper into the foundations. What makes the fact that you are reading this book unquestionably true (capital T)? Here, most people come up with a long list of seemingly infallible ways of establishing this. Well, you might say, I can see the book (or device where you're reading it), touch it, smell it, I can even hit and doubter with it!

Maybe, right about now, you're beginning to feel a bit exasperated with this (ridiculous?) line of questioning. Please don't. Remember: inspecting foundations is dangerous and scary. But wouldn't you agree with the fact that such exploration is of vital importance. After all, if facts are not what they feel (let's go back to the scenario of the evil genius), our entire conception of the universe collapses. If you're still not convinced of the value, I say, the utter mortal-danger necessity of asking such questions, let me show you with an example.

Right now, as I'm writing these lines, elections in the United States of America are being, in large part, determined by unexamined answers to knowledge questions like the ones above. Is climate change a fact? Is the claim that humans are the cause of it a fact? How do we know any of this? In the heat of the moment, most of us answer these questions by political affiliation. But, with a cool head, we all know that that fact-finding recipe cannot possibly work. As John Adams put it "Facts are stubborn things; and whatever may be our wishes, our inclinations, or the dictates of our passion, they cannot alter the state of facts and evidence."[1] In this case, and without any exaggeration, if the science is right, following our inclinations instead in the case of human-caused climate change might cost us our human-friendly planet.

Foundational questions are painful to ask, but their questioning is not optional. The alternative is to follow our guts, not to question them at all, and to keep going as if nothing is happening. But as our illustration above indicates, picking quick and dirty, or arbitrary answers to foundational questions can turn, sooner than later, catastrophic. So, dear reader, in this book, we will dive into the questions that make up the core of the question of fact-finding. Most likely, it will not be fun (foundation digging is bound to be painful), but it is essential, and hopefully, it will also be rewarding – at a deeper intellectual level.

---

[1] Reference

1.2 *Thought Experiments*

As hinted above, asking foundational questions tends to make people uncomfortable, and to put them in a defensive, core-belief protection mode. In fact, we seem to be designed to have several layers of hard-held beliefs, and a number of effective core-defense mechanisms, that shield foundational issues from pesky lines of questioning. But there's an effective way to breach these defenses, the careful contriving of *thought experiments*.

A thought experiment, like any other experiment, is designed to bring up or de-clutter a particular variable that would otherwise, in a realistic scenario, be buried – and, consequently, would not be understood, would remain mysterious, or even go unnoticed. In a thought experiment, however, the experimental setting does not consist of a lab with a set of apparatuses: it consists of an imaginary scenario. As hinted above, in an actual experiment, most of the apparatuses are designed to isolate the variable in question from other confounding variables. In a thought experiment, however, the same isolation is accomplished through crafting the right set design around it. So, be warned: the imaginary scenarios of a thought experiment can get to be quite outlandish. But remember that the point of a thought experiment is not to make you think that the outlandish scenario is believable, or even possible, but to strip a target question of all the day-to-day circumstances that otherwise would make it difficult to address and assess and that would sometimes, even, render the question invisible.

Questioning the idea of a *fact* without the help of a thought experiment could not only be annoying but, more importantly, very difficult. So let me tell you one such useful, outlandish scenario that will, like any good thought experiment, help us to see some difficult questions and will disclose some unseen angles.

1.3 *The Seed Ship Malfunction*

In 2098 (a time that later came to be known as "The Great Collapse"), the Anthropocene era (the age of humans) ended. By the later half of the XXI century, global climate change hit the planet with a vengeance -- much, much faster than we had ever thought it would. And as many scientists had warned, the human global system, too fragile and convoluted to bounce back, collapsed. Wars, famines, and pandemics wiped most of us out in an astonishingly short period of time. But some of us left before it was all over -- sort of. Before The Great Collapse, a team of scientists working under the mandate of the United Nations feverishly put together a desperate plan to ensure the future of our species. They created twelve Seed Ships that would flee our soon-to-be-inhospitable Earth, each one in a different direction, aimed to what we hoped would be a new, human-friendly home. The ships were loaded with extensive robotic libraries, factories (including bio-mechanical wombs), and frozen-embryo collections. The collection included, of course, human embryos. And this is where your thought-experiment self comes in.

Hundreds of years have passed and you are one of the 150 human embryos that were carried to term on Gliese 581g (an Earth-like planet discovered in the early 2000's).[2] You live in the complex of automated farms, factories, and living quarters that sprouted out of Seed Ship #7. Seven was the ship aimed to Gliese 581g, but that ship is now unrecognizable. The original vessel sprouted a veritable village. Robotic factories produced all the necessary machines to ensure that the planet was earth-friendly by the time the human embryos were carried to term. Specialized robots were also supposed to be ready to become your nannies and teachers, but unfortunately, they were severely compromised early on in the terra-forming process. Surely, even if everything would have worked the way it was supposed to work, we knew you were going to turn out a bit weird. After all, no one had any reason to believe that our (still) crude robotic intelligences could replace a nurturing community of humans. But without the nursery- and teaching-robots working to full specs, your generation's batch turned out even weirder than we had expected. You were all well fed, and you developed some rudimentary language skills, but without access to our hard-won, carefully- accumulated Earth knowledge, every member of the Gliesean human community was (to put it mildly) severely challenged. So, now, let us fast forward a bit: imagine that all Gliesean humans are in their early teens, and let us put our thought experiment to work.

---

[2] O learn more about this *real* planet, you can go, for instance to NPR's *Science Friday's* website: http://www.sciencefriday.com/program/archives/201010013.

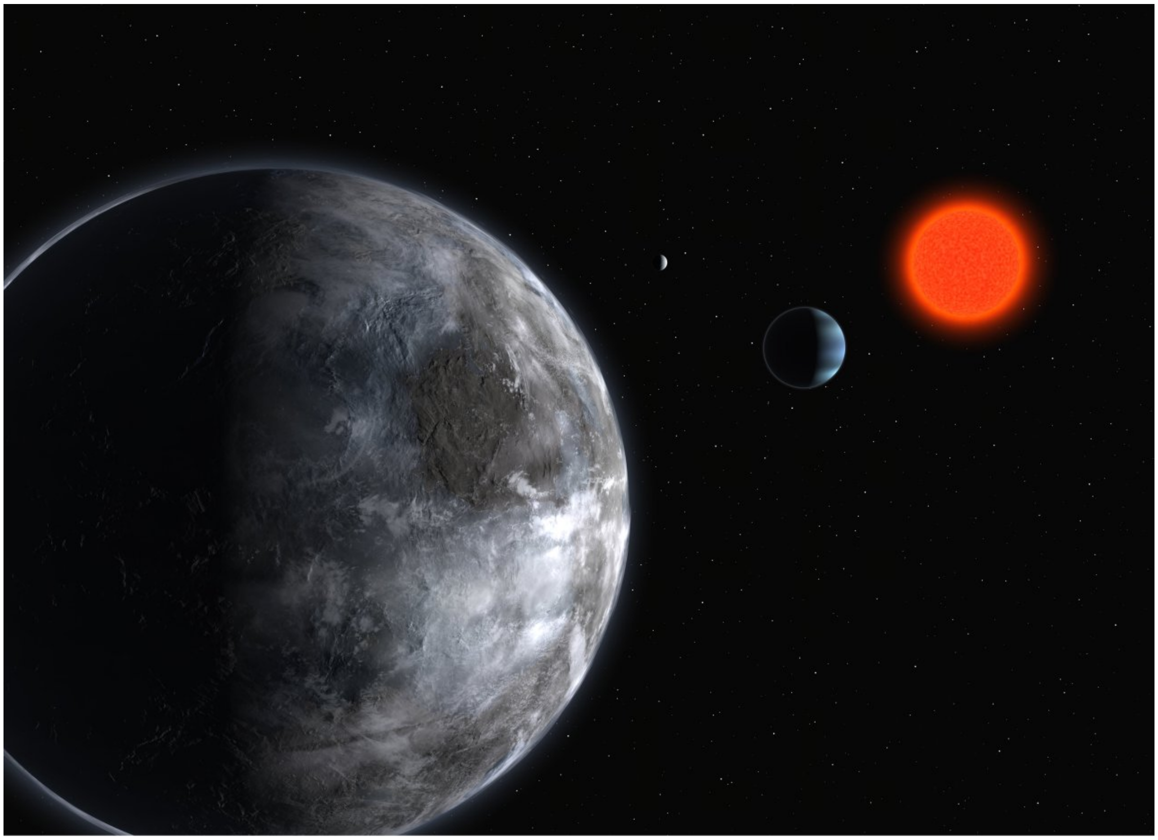


Figure 1-1. Gliese 581g is a real, Earth-like planet that was discovered not too long ago. Here is an artistic depiction of it, produced by the European Southern Observatory and posted for common use at http://commons.wikimedia.org/wiki/File:ESO_The_Planetary_System_in_Gliese_581_(by).jpg.

I hope that you are already asking yourself all sorts of interesting questions. For instance, would the Gliesean language have the same ideas behind it as our Earth-born languages? Or, to put it slightly differently, would Gliesean be as translatable into English as, let us say, Japanese, or Spanish is? Would it have fewer or more words? And if so, would there be a pattern behind which words are most likely to be missing, or in excess? Would Glieseans have concepts for cause and effect; or for time; or for property? Whether you realize it or not, the thought experiment is already working. This scenario has made us think about issues that we would have otherwise rarely considered. We are already asking all sorts of foundational questions about the interaction between the world and our minds. For at least the last 2000 years, philosophers and (more recently) cognitive scientists have spent gallons of ink dealing with questions that are very similar to the ones we have just brought up. But now we must rein in our thought experiment. The problem with foundational questions is that they come in layers, like an infinitely deep onion, in which the innermost core is forever out of reach: there is always a deeper layer of questions that one must grant in order to proceed. So, you arm yourself with patience and temporarily, for the sake of the argument, accept the layers below, establish an initial foothold, and start your critical analysis with the layer at hand. For the sake of our present purposes, I would like us to concentrate on the layer of questions that deal with the problem of how Glieseans know about their world. Or, to put it into other terms, how do Glieseans find their facts.

The first stumbling block we should encounter in our quest is that of clarity about what we are after: we must be clear about what we mean by "fact." But this word is surprisingly elusive and, in my experience, as soon as you question it, four other words (even more elusive than the first) pop up to "help:" objective and subjective, true and false. So let us be clear, at least, about what we usually intend this family of related words to mean; let us start first, with the facts.

1.4 Facts

When we say "fact" we are usually talking about those beliefs about the world that we hold as certain. So we know for a fact that the sun comes out each morning (albeit, sometimes behind the clouds); that things fall when we let go of them; that we are not dreaming right now; that the Earth is round; and so on and so forth. "Fact" is the word we reserve for those about-the-world beliefs to which we attach certainty, for which doubt to even be considered seems silly.

It is important to realize that when we ask "what is the shape of planet Earth?" versus "what is the best flavor of ice cream?" we are asking two very different kinds of questions. The first one is about planet Earth, while the second one is about us. We highlight this vital distinction by saying that the first kind of question is about matters of fact, and the second kind is about matters of opinion. Matters of fact claim to deal with the outside world while matters of opinion with personal preferences, about the world or otherwise. Naturally then, it would be silly to try to argue that, as a matter of fact, vanilla is the best flavor of ice cream. It would be equally silly to argue that the shape of the Earth is a matter of opinion. So, for the purpose of our present discussion, we can say that the distinction between matters of fact and matters of opinion hinges on the subject matter we are targeting. When we are trying to say something about the state of the outside world, independently of the subject that is doing the talking, we are dealing with matters of fact. When we are trying to say something that involves the subject who is doing the talking, we then enter into the realm of matters of opinion.

Before we go any further, I must stress that this distinction (between matters of fact and matters of opinion) is of utmost gravity. Fact-finding (or "not-finding") could really be deadly -- and not only for Glieseans. To internalize this, we might need to propose another short thought experiment. Imagine if I could, somehow, make all the "facts" in your head disappear. What would happen next? Would you be able to remain seated in the chair that you (factually) trust; expect not to float up (which you factually know to be impossible); trust your legs to stand up; your mind to conclude what to do next; or the conviction you are not mad? With your facts magically removed, the most profound sort of uncertainty possible about the world (including yourself) would paralyze you. This should suffice to establish that facts are essential for functioning; being able to hold some of our ideas as factual is necessary. Or, to put it slightly differently, we need facts in order to act. Now that the gravity of the issue is on the table, and in order to keep going, we need to tighten the distinction between matters of fact versus matters of opinion by dispelling a possible complication.

At this point, you might be pondering the following question. What about those times when, for instance --a subject-- talks about another subject (or even myself!) as an object? For example, what about those times when I want to talk about a subject's (myself or another's) brain states when, let us say, ice cream is tasted? Not to worry, this is only superficially confusing because, in a case such as this one, you are still talking about an object (which, in this case, just so happens to be simultaneously capable of being a subject). The crucial issue here is whether the statement in question (the one we are trying to decide as a matter of opinion or a matter of fact) is mostly dependent on the person doing the talking. Or, to put it slightly differently, matters of opinion can change as the subject doing the talking changes, while matters of fact cannot. In our example, statements about the state of our brains should not change depending on which subject is experiencing them, or talking about them. So even though it might sound a bit counterintuitive, and as this example illustrates, we can talk about the flavor of ice cream from the matters-of-fact angle (like when we ask questions about the brain states behind the issue) as well as from the matters-of-opinion angle (like when we claim that vanilla "rules").

1.4.1 Objective versus Subjective

Now, when we say *fact* not only are we claiming that we believe something about the world, but also that anybody in their right mind should believe it also. This angle of universality that we expect facts [are intended] to have is captured by the related ideas of objectivity and subjectivity. For the purpose of our present discussion, we can say that when we talk about matters of fact, objectivity usually refers to the stance and approach that one adopts in order to obtain them. In other words, when we (a subject) try to say something about an object, and do our best to extricate ourselves from the picture while doing so, then we are being objective. In the same vein, when we use methods that ensure the latter, we also say that those methods are objective. As I pointed out before (in the case of matters of facts versus matters of opinion), when a subject succeeds in being objective, then what (s)he says shouldn't change if the subject changes. To give an example, if Jack says something objective about the shape of the Earth, then Jill should have no choice but to say the same thing about it. Thus, when our stance is objective we maintain that any subject should get the same facts as long as they adopt the same stance and follow the same procedure. Now we are ready for true versus false.

1.4.2 True versus False

Let us recap. We understand that when we are trying to talk about matters of fact we are also usually trying to be objective.[3] Now the added layer of discussion is that when we are trying to be objective, there are two possible outcomes: we either succeed or we do not. We need a new set of words that helps us distinguish these two possible outcomes. We can say that a statement about the world out there is true when what the statement *says* matches the way the world *is*, and false when it does not. So, all together now: when we talk about the world and try to be objective we adopt a particular (objective) stance and procedure that produce factual statements; when those factual statements match the world, we say that they are *true,* and when they do not, we say they are *false*. An interesting outcome of our discussion is that we can objectively produce a fact, show it to be

---

[3] I will not discuss here those cases where one is trying to establish a fact via subjective terms. Think, for instance, of those that attempt to establish the factual existence of a supernatural being by personal revelation. As this lies outside the scope of critical analysis it also lays outside the scope of the present treatment. But see more of this under the heading "**How to Know: The Recipe.**"

true, but later on realize that it was false. Or, to put it slightly differently, objectivity does not guarantee truth.

You might feel, at this point, that we have detangled the ideas of objectivity, fact, and truth to the point of confusing ourselves. Some might think that if I grant that a fact can reached objectively then later be shown to be false, then I must surely be tripping over my own tongue! But thinking this would most certainly be mistaken, so let me explain why with the aid of an example. Let us say that you have every good (objective) reason to conclude that there's a hungry tiger in the room beside you. You cannot see it, but you heard of the man eating tiger that escaped the circus an hour ago, a couple of blocks away; you can hear the tiger; you can even smell the tiger (you happen to be a professional tiger biologist). You can also hear that the tiger is slowly but surely gnawing a whole through the door. Decision-time: should you leave your six-month-old baby sister in that room alone while you go upstairs to take a nap?

Of course you would not. But why is that the case? Is it because you know that it is absolutely true that there is a man-eating tiger on the other room? Well, not exactly: you have no incontrovertible proof; you do not really know that to be a fact for sure. Obviously, the question of certainty (or lack thereof) is not what matters here. What matters is that you have tons of evidence pointing to a tiger. What matters is that you have many different lines of objective reasons to grant the tiger factual status even though, in the end, it might all turn out to be some (obscene) prank. In other words, the question of whether or not the would-be fact will end up being true, or not, is not that relevant when the task is to discover a fact and determine the appropriate course of action.

So, already there is a rather counter-intuitive lesson to take home: objectivity, not truth, should be the tipping point when one considers whether or not to take a would-be factual statement seriously. Or, to put it slightly differently, objectivity --not truth-- is what *should* be the matter when we decide to bump a statement about the world from the "would-be-fact" category to the "full-blown fact" category. The objectivity behind the statement in question (the one ingrained in the procedures we followed in order to conclude that there was a hungry tiger gnawing at the door) puts us in a situation of moral responsibility *regardless* of whether or not the statement turns out to be true or false.

Furthermore, if one thinks about this issue a bit more, one should soon realize that the expectation that our humble brains can possibly capture a perfectly truthful conception of what is out there is, to say the least, presumptuous. Well, let us say it as it is (even though it hurts): it is actually a bit silly. There is no reason whatsoever to expect that the universe must fit the biased conceptual limitations of one species of primate that happens to have arisen in some insignificant corner of some insignificant galaxy.[4] We must bite the bullet that our humble origins make inescapable: truth is a nice ideal that, when examined, up close and personal, crumbles as nonsensical and thus unattainable. A more realistic ideal is to aim for those facts that we can reach after being as objective as one could hope of this humble species of primate. We know that these facts will not be pure or perfect conceptions of the world. But they will certainly apply and hold for any subject that also tries to extricate its idiosyncrasies as far as (s)he possibly can. That is as close as the truth as we will ever get.

So, to sum up, we first realized that in order to act, we need facts, but now we understand also that the way we get our facts --and not only whether or not they end up being true or false-- is also binding. This is important, so let me say it again, but slightly differently: we should take our facts seriously, primarily because of the way we reach them (objectively versus subjectively), and not because we know, for sure, that they are true. In fact, we will realize soon why the latter cannot even happen! We will not ever know for sure that our facts are true; we can only feel assured in the knowing that the way that we reached said "fact" was objective.

1.4.3 Gliesean Fact-Finding

The crucial question that we are now ready to ask is how do we know we're being objective. In other words, how do we know we have reached that point where we should accept the concluded fact as if it were true (even if, later on, it turns out to be false). To tackle this, we need to dive deeper into our thought experiment.

We now understand that in order to act, first we need to have a set of facts that we have established; this, we must allow Glieseans some facts[5].

---

[4] If you have a problem being a primate, please hold your discomfort until you get to see the demonstration later on in this essay.

[5] This is quite unrealistic; you should ask the perfectly legitimate question of where on earth did these facts come from. But please remember that thought experiments must be reined in and that the point here is not to argue that this

Let then our Glieseans know for a fact all the elements that are necessary to ponder: let us give them the most minimal but necessary facts to step out of their (otherwise unavoidable) mental paralysis. For instance, we will allow them to know for a fact that their senses are mostly reliable; that they are not in a dream; and that some of the ideas that their heads can concoct are not true. For instance, they know for a fact that they can imagine themselves flying but that they cannot actually fly. In other words, they understand that the statement "I can fly" is false and that "I cannot fly" is true. Let them also have a minimalistic set of facts about their immediate surroundings. All in all, the teen Glieseans find their heads full of would-be facts about themselves and their immediate surroundings, but know nothing about their new planet and, most importantly, know nothing about how to find new facts. Glieseans know nothing about "how to know."

1.4.3.1 How to Know

Put yourself in the shoes of our unfortunate descendants and wonder what would you say if, in a (desperate) community meeting, you were asked to propose a recipe for knowing, for getting new facts. Your first conundrum would be to decide what kind of proposed recipes the community of brand-new knowers might consider acceptable.

1.4.3.2 I Feel I Know versus *I Know*

The first problem Glieseans will have is that everybody thinks they know the facts and that many are ready to insist on it. So, what to do with the intensity of the would-be knowers' beliefs? Imagine the following exchange at the community meeting:

> Jill: "I *really, really,* know this. I *feel* it with all my heart that giant Gliesean tarantulas would make wonderful pets."
>
> Sean: "Gliesean tarantulas have paralyzed and eaten every single one of us that has tried to pet them. I believe for a fact that they are dangerous."

Is the strength of one's beliefs any indication of their accuracy? To answer this, one must concentrate on the subject matter in question. If we were inquiring about the subjects themselves, then their beliefs would, in fact, be relevant. But here we are asking questions about the outside world (and in it, a giant, tarantula-like creature, to be precise). So ideally what we are

---

scenario is possible, but to create a setting that helps us detangle issues that everyday-life makes otherwise too complicated to tackle.

aiming for are statements about Gliesean tarantulas that contains as little from the subject doing the knowing as humanly possible. Ideally, the content of these statements should be all about the world and none about us. Because of this goal, and though we might care deeply for Jill and Sean, what they feel is not really relevant. To put it briefly, we want facts to be objective. So, the first decision our young Glieseans would have to make is that intensity of believing says little to nothing about the objectivity of a would-be factual statement. Hopefully the objective force of a fact does make us believe it intensely, but the reverse is not true: our intense believing does not turn our beliefs into facts. So, here comes the big question: if the intensity of believing does not tell us much, how then can we know that a fact is a fact?

1.4.3.2.1 Facts: Processed Results or Passively Perceived?

Are facts something that we perceive directly, or are they the result of careful mental processing of other information? You might think so. After all, if I say "It is a fact that this page, or screen, is in front of you right now," it should surely feel like you know that this is true, directly and instantaneously. But this feeling is misleading. The fact that there is a page in front of you is not taken in passively by your mind; it is a conclusion that has been, taken in, processed, and crunched out by it. This fact, like all others, has been processed or reasoned. It is a conclusion you have reached given a particular set of elements you are presented with. These elements, most usually and most prominently, are your perceptions combined with relevant bits of past experience about the piece of the world in question. Let us rewind now, so that we can better dissect and understand this false sense of instantaneity with which facts present themselves in our minds.

What you actually did when you seemingly instantaneously decided that there was a screen or page in front of you was to move from your trusted old perceptions of screens or pages, to the conclusion that there was one in front of you. Though it might not feel like one, the unveiling of this lightning-fast processing is a momentous discovery, so we should pause and take it in slowly: we have just realized that facts only *seem* to appear to us *passively* – or without any processing on our parts. But in truth, facts are the result of mental processing's and maneuverings that go from a set of reasons (that we will from now on call "premises") to a given conclusion (that is, the fact itself). Let us look a bit more carefully at all the premises in this lightning-fast piece of thinking.

Amongst the premises that your mind assessed, it should be easy to see that your perceptions figured most prominently (but were gathered and accepted so quickly that you did not even notice). By *perceptions* I mean such things as the input from your eyes and all other senses (like touch or smell) to the brain. Of course, there were other premises included in your super-fast thinking, like your past experiences (with pages, screens and such), as well as your general understanding of the present situation. All in all, you had been exposed to similar sets of premises before, and you had gotten very good at automatically concluding that a set like the one at hand now, 99.9% of the time, was best explained and thus meant that there was a real page or screen in front of you (and not, let's say, a hallucinated or imagined one).[6] Now let us apply this realization to the Gliesean conundrum.

1.4.3.3 Assessing Would-Be Facts

After initial discussion and careful considerations, Glieseans should conclude three very important things:

[6] Maybe you are now thinking that, in our dreams, 99.9% of the time, the same lightning-fast reasoning fails to alert you that you are only dreaming. That is, most of the time, in one's dreams, one concludes, *with the same lightning-fast piece of reasoning we have just outlined,* that everything in the dream is factual. This is a very telling and profound piece of thinking, but we are not ready to tackle it just yet, so please bear with me. I promise that we will deal with this shortly.

1- Facts Must be Taken Seriously – Regardless of Whether or Not They End Up Being True or False
Our facts might turn out to be true or to be false. But once we find one, by adopting the most objective stance possible, it is our responsibility to take them seriously. (We saw that those that tend to disagree with this statement might also tend to find their young relatives eaten by hungry beasties in adjacent rooms.)
2- The Strength of Belief is Irrelevant
Finding statements about the world by adopting an objective stance should boost your confidence in the statements under question. But no amount of inflated confidence, no amount of hard believing, will turn a would-be fact into a fact. What turns would-be facts into facts is good thinking. Consequently:
3- To Assess a Fact is to Assess the Methodology Used to Reach That Fact
Since facts are not taken in passively but are the result of objective thinking, in order to assess a given fact one must assess the thinking or method used to obtain or conclude it.

Finally, Glieseans have a target of study. They should now understand that figuring out "what is the best recipe for fact-finding" is one and the same as assessing "what is the most objective methodology." Considering how much subjects believe the would-be facts in question, or whether the would-be facts will turn out to be true or false, are side questions. Now, the Glieseans and we are ready to start looking for the most objective recipes out there.

1.4.3.4 The Recipes

So what are the recipes for knowing, for getting new facts? After thinking for a while, most Glieseans (just like most of us Earthlings) would propose recipes that include a certain mix of four ingredients, which I will call the trial and error method, the best-explanation method, the pure-reason method and the categories method.

1.4.3.4.1 The Trial & Error Method

First, we propose and suppose that some fact or other is the case. Second, we think of a particular state of affairs that could only be the case if and only if we were correct in our supposition. Third, we look for this state of affairs, or try to bring it about. Four and last, if we succeed in step three, we conclude that our original proposition was correct, and if we do not, we start all over again with a new proposition. Let us try it out with the Gliesean tarantulas we considered before (please do not try this at home, not even with Earth tarantulas).

1) We propose that, if Gliesean tarantulas make good pets, then the state of affairs that they are nice and cuddly should be the case.
2) We send one member of the fact-finding recipe committee to bring about this state of affairs (or lack thereof) by petting one of them.
3) The tarantula eats the committee member.
4) We conclude that Gliesean tarantulas do not make good pets.

1.4.3.4.2 The Best Explanation Method

Let us go back to a more mundane (and less dangerous) fact that we also discussed earlier, the one about the existence of this very screen, or piece of paper, in front of you now. If I were to request the most immediate list of reasons you used to conclude its existence as fact, a list of *perceptions of the screen/page* would diligently pop into your mind as the main justification for concluding the *actual screen/page*'s existence. Typically one would conclude the screen/page's existence when one can see it, touch it, even smell its peculiar smell, and/or hear the noise one's fingers make when they touch it. But, you might ask, very perspicaciously: how do we go from a list of perceptions of a screen/page to the existence of the screen/page itself? Well, if you were very adventurous and were to go around asking your friends this very (annoying) question, I bet you they would respond something along the lines of: "Well, what else am I supposed to conclude? That I am in some kind of virtual reality? That I am hallucinating all these perceptions? *Obviously*, the most *obvious* explanation from all my perceptions of this screen/page is that *there is* a screen/page in front of me!" (Incidentally, I would recommend that you abandon this line of inquiry if you want to keep your friends.)

1.4.3.4.3 The Pure-Reason Method

Another way of reaching the state of certainty that we associate with the idea of facts has little to do with trial and errors, or our best explanations and, for lack of a better title, I call it the pure-reason method. It would be hard to argue that you know, for a fact, that "8 divided by 2 equals 4" has anything to do with our perceptions of 8's; 2's, and 4's (if you doubt this, run the same idea with 364's; 2's; and 182's, or alternatively, with the idea of the square root of 3). These pieces of reasoning are about abstract, as opposed to concrete, ideas--meaning, for instance, that you cannot know them by manipulating actual (i.e. concrete) objects. The number 767 or the letter α are mental constructs, symbols one can invent to explore relations such as *included in*; or *greater than*; or *necessarily followed by* – to mention just a few. It should be easy to see that one does not need to get intimate with ε to be certain or know for a fact that if all X's have a certain property €, and ε is an X, then ε has property €. You know this *not* because you have met many X's and €'s, but because you understand what a symbol is and what this abstract relationship entails (that of *bearing a property*).

We certainly apply these abstract moves (very usefully) to the world of concrete objects. We can divide a flock of 40 sheep into equal portions, but have to do more abstract maneuvering when we have a flock of 41. Similarly, once we have discovered that all mammals have some hair, and that dolphins are mammals, we can say that dolphins have some hair. We can even predict the pattern of growth in a head of broccoli once we understand the abstract formula behind its development and growth.

1.4.3.4.4 The Categories Method

This is the very useful method by which we come up with categories and sweeping generalizations. We collect (so to speak) instances of a particular kind of event until we feel confident that a pattern is clearly emerging. Then we conclude that whatever it is that we found in all of the particular instances we sampled is also the case for every member of the category in general. Imagine, for instance, that a particularly slow member in our fact-finding recipe committee insists that only one (albeit very unfortunate) instance of tarantula misbehavior does not amount to a good reason for making a pattern or sweeping generalization about *all* Gliesean tarantulas. After careful consideration, we decide that a second trial-and-error tryout should be performed. We are thus forced, to our great sorrow, to record our second instance of tarantula misbehavior. But we are thankful to our second fallen committee member for demonstrating the value of careful repetition and the connection between a larger sample size and the confidence we have in the resulting generalization. So, accordingly and to honor her memory, we wait weeks and weeks before we make other generalizations --that the sun comes up every morning, for example-- and we only compromise our eagerness to obtain forever-increasing sample sizes when dealing with dangerous categories (such as any related to committee-member-eating local fauna).

All in all, and after thinking about this long and hard, most Glieseans (just like most Earthlings) decide that all factual thinking involves maneuvers that can be broken down into different combinations and permutations of these four methods. But what have they accomplished thus far? One could say that they have gotten a grip on *how* we do it. But the million-dollar question still remains: we certainly trust these methods, but how can we tell if they are as good as we feel they are?

1.5 Analyzing the Methods

Since we have discovered that facts are the result of one piece of reasoning or other (what we called above *recipes*), our analysis of these recipes must necessarily boil down to an analysis of pieces of reasoning. So, to analyze the fact-finding recipes we will have to break them up into their components: premises and conclusion/s. Unfortunately, as you will see, this is easier said than done. So at this point it is necessary to introduce a formal tool to help us do this. It is called "argument reconstruction."

1.5.1 Arguments and Reconstructions

From now on, pieces of reasoning that include a set of premises supposed to lead to a conclusion will be called "arguments."[7] In order to study them, we will present them in a way that emphasizes the backbone they are resting on: bare-bone premises and conclusion. I say "bare-bone" because we will try to present an argument's premises and conclusion in the most succinct way possible while still preserving whatever force makes the conclusion follow from its set of premises.

Reconstructing arguments is very, very hard. We tend to be (unduly) impressed with our own thinking and consequently, it's very difficult to expose that thinking and make it vulnerable in such a way. But reconstructing arguments is a necessary part of exposing the (supposed) connection between premises and conclusion so that we can objectively accept or reject what the piece of reasoning is asking us to do. Remember: an objective piece of thinking remains so regardless of whom is doing the

---

[7] Please do not be mislead by the colloquial use of the word *argument*. Here, we will use it as a technical term. Here, this word does not mean that any kind of animosity is entailed or should be assumed. On the contrary, the whole point of treating premises and conclusions formally, like arguments, is to assess their objectivity.

thinking. So once one discovers that a given piece of reasoning is objective, then that person, as well as everybody else, is bound by it – regardless of whether one likes it, or wishes to believe in it. Reconstructing one's own arguments is also particularly helpful since it allows us to get rid of the noise in our thinking. Noise in thinking can make you drop facts that you should not have dropped and keep would-be facts that you should have dropped. As we learned above, this could be quite deadly.

1.5.1.1 The Concrete and the Abstract

When we reconstruct and then critically assess an argument, we concentrate on two important aspects of it: one is its content, or what the piece of reasoning is actually claiming about the world; the other is the abstract, or relational, operations --as opposed to concrete operations-- amongst the elements that bear content. Let me explain this with the help of an example. Let us start by reasoning that:

All dogs have some hair,
Fido is a dog,
then Fido has some hair.

1.5.1.1.1 The Realm of the Concrete: True or False

The first aspect, what we loosely called "content," is analyzed when one considers whether or not the premises are true or false (as described above, in the sense of their being a match, or not, between the statement and the world). But right now, we are not going to worry too much about this aspect. Instead, we will concentrate on the degree of strength with which the conclusion follows from these premises (there will be more--much more-- on this to follow). In this piece of thinking, the content--that is, what we are saying about the world of dogs, of Fido, and of the property of hair are all true. Checking whether or not this content is true--that is, checking the truth-value of a premise, to use a more technical term, -- is not always this easy. But we will discuss much more of this later, and so, for the present example, we will say that the content gets its check–mark of approval.

1.5.1.1.2 The Realm of the Abstract: What Stands to Reason and What Does Not

Assessing the relational operations that an argument involves can get very technical but, for our present purposes, we do not need to worry about that yet. At this stage, it should suffice to realize that arguments do more than (ideally) present true premises and conclusions: arguments perform relational moves amongst the elements in their premises so that the conclusion follows from them. This should be easier to grasp with the help of our last example, and the first step is to discover the elements whose relations we are going to deploy. In our example, these are dogs in general, Fido in particular, as well as the property of “some” hairiness. Now, let us consider the relational maneuvers to which we are subjecting these elements. We are saying that, if a given kind (i.e. dogs) has members (e.g. Fido), then, the key properties of the kind (i.e. some hair -- just so you know: all mammals have some hair and dogs are mammals) are also the properties of each member. The relational operation that we are using here is that of inclusion, or membership, to give it a name. A quick analysis should deem that it is correct to say that if there is a kind, then all of its members will share the properties that makes one a member of that kind. Please note that this relational operation between member of a kind and kind itself is not reversible, i.e. it is not correct to reason the inverse of the above: that key properties of a given member must also belong to all other members (e.g. not all members should have the name “Fido”). It is also important to note that while we analyze content by checking whether or not it matches the world, we do not need to inspect the world to see if the relational operations in a given argument stand to reason. You might be wondering how it is, then, that we know what stands to reason in terms of relational operations. This is an excellent question that has been bugging philosophers for over two millennia. There are some excellent answers out there, but exploring them would takes us well away from the scope of this piece. Let us agree then, for the sake of the present discussion, that most humans seem to come with the necessary software to spot what stands to reason and what does not.

Now, for the sake of clarity, let us agree on a standard way of representing arguments. We will identify premises with the letter “P” and conclusion with the letter “C.” We will separate premises from conclusion by drawing a line between them. If there is more than one of each in an argument, then we will add a number to each one of them and we will start with the most general premise or conclusion and move towards the less general. So, in the case of Fido’s argument:

P1: All mammals have some hair
P2: Dogs are mammals
P3: Fido is a dog
C: Fido has some hair

1.6 The Art of Argument Reconstruction

Now that you know what ideal arguments are made of and what one is supposed to assess in them, we need to learn to identify their wilder, not-so-ideal counterparts: the arguments we use in daily life, those that you are likely to encounter when reading the news, or talking to friends. These arguments of the wild, even the good ones, rarely come with clearly-outlined and structured premises and conclusions. So it is important that you know how to extract these elements in order to assess the arguments' strength. Let us try to do this with the help of another simple example.

I hope that you know for a fact that every person around you is mortal. Anyone trying to present their reasoning behind this fact might end up producing something like this:

> Well, since the beginning of modern history we have known that every single person has died. Besides, we are all mortal, so we all must die. Also, every doctor understands why we are mere mortals. So everyone around me, most positively, must be mortal.

As you re-read this stream of consciousness, you should be able to see that there are many elements that could be considered premises, and a few that could be considered conclusions. Usually, the former are associated with an explicit- or implicit- *because*, while the latter follow a *so*, *therefore*, or other conclusion that is also either explicit or implicit. In our example, both of these circumstances are the case. We know that "everybody around you is most likely mortal" is the conclusion both due to prompt and key word (for example "so" in the last sentence).

The next hurdle is to identify the argument(s)' *structure*. Many times we argue with no structure in mind whatsoever. In these cases, what we most commonly do is provide a set of premises that are *vaguely* related to the conclusion. Technically speaking, we are not *arguing* a conclusion; we will call this *sandbagging* a conclusion. Putting it more precisely, we sandbag a conclusion when the latter does not *follow* from the premises; there is no relation amongst the elements in the premises that inevitably, or with some degree of likelihood, leads to that conclusion. Sandbagging makes you think about the conclusion but it does not *force* you, objectively, to arrive to it. Let me give you two examples.

The conclusion *follows* inevitably from the premises because they entailed it. We are not sandbagging.	The conclusion *feels* right, but the premises do not entail it, or make it follow with any force. We are *sandbagging*.
P1: A is included in B P2: B is included in C C: A is included in C	P1: It's late P2: I'm tired P3: I don't even like Wilma C: I'm not going with her to the movies

Why is this important, you may ask? Why should you be able to distinguish between *sandbagging* a conclusion and *arguing* one? Remember, facts should be the result of objective thinking. When you *argue* a fact, it is being *derived objectively* from accepted premises. In other words, the content and relational properties in the premises *determine* what the conclusion should be. When a conclusion is sandbagged, the subject's feelings determine whether or not the conclusion is accepted. Now we can spot arguments among sandbags, so let us go back to the original stream of consciousness.

The stream of thinking above seems to provide at least three "because's:"

In the past everybody died.
We are mortal.
We understand that dying is an unavoidable part of the process of living.

It would be quite tempting to reconstruct this as a single argument, with three premises and one conclusion. But that would be sandbagging. The three premises are not working *together*. Or, to put it more precisely, they are not different elements in some mental formula, so that the conclusion follows *only if* we articulate the three elements in some precise fashion. These are three independent arguments:

P 1 All humans have died in the past
All humans have and will always die

P1 Life process 1
P2 Life process 2
P3 Life process 3
PN Life process N[8]
Death best explains the above set of processes

P1: If a given life form is mortal, then it must die
P2: Humans are mortal
C: Humans must die

Now that the arguments are reconstructed, we are finally ready to learn the skill of assessing them. But that will be our job in the next chapter.

1.7 A Summary of What We have Learned

Let us review what we have accomplished. We enlisted the help of a thought experiment in order to consider a question that is usually buried under too many unquestionable aspects of daily life: how we arrive to our most trusted, blindingly-obvious facts. We realized that the idea of *fact*s is tangled up with the related ideas of *subjective* versus *objective*; and *true* versus *false*. We understood that the first word pair refers to the different stances that one can take when talking about the world. One is *subjective* when (s)he is mainly talking about the way the world *feels* to him/her and *objective* when (s)he is doing everything humanly possible to extract him/herself from the picture. Regarding true versus false, we understood this dichotomy to refer to the relationship between a *statement about the world*, and the *world itself*. When the statement *matches* the world, we say the statement is *true*, and we say that it is *false* when the statement does not match the world. Surprisingly, we concluded that the objective nature of our fact- finding methodology was what made the resulting facts binding, *not* whether or not they turned out to be true. Consequently, we discovered that most important question to ask regarding fact-finding was how do we know that the methodology we used was truly objective (and not, for instance, how strongly we *believed* in the fact under question).

[8] For us, the capital letter N indicates some unspecified very large number. In this case, it indicates that there are many, many processes that are explained by the idea of *death*.

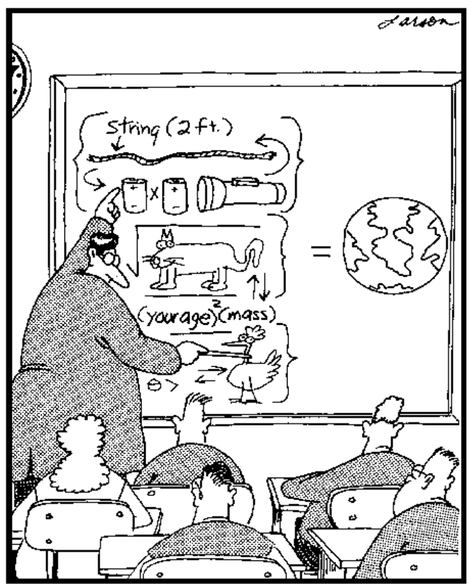
Facts can be deadly, since they are the necessary bedrock behind every action we take. We also discovered that they do not passively appear in our minds, but are actively pursued and result from our thinking: facts are the conclusions of lightning-fast, usually-unnoticed arguments (of which the most commonly-used premises come from perception). We saw that arguments should be checked for their content (what they are concretely saying; their true/false value) as well as the abstract relationships amongst the elements they contain in them. In order to assess these vital pieces of reasoning, we learnt the first necessary step: how to reconstruct arguments so that we clearly disclose the elements in their premises and conclusion, as well as all formal relationships amongst those elements.

As I mentioned above, now comes the final and [probably] most critical part. With the argument cleanly reconstructed, one now needs to decide if it is a good one or not. Remember that we will deal with the content of the premises (or the true versus false value of their elements) later. What remains is at least as difficult to assess. More precisely, what we will ask now is whether or not the conclusion really follows from the premises.

2 CHAPTER 2 -- CONCLUDING FACTS VIA DEDUCTION

2.1 Does it Follow or Does it Not -- That is the Question

Look at the cartoon below. Most people find it funny and now I will ask you to wonder why – just for a few seconds, before you continue reading. Okay, now, let me venture an explanation: most people find this cartoon funny because it ridicules the very question we are trying to explore: when a conclusion does or does not follow from a given set of premises. I would say that this cartoon is funny because we do not have to think hard in order to realize that batteries, two feet of string, and chickens cannot possibly have anything to do with the creation of our planet. It is funny because we instantaneously know that the conclusion, in this case, does not follow from the premises. The following should be interesting to note: this cartoon does demonstrate that we have a very good grip on whether or not a conclusion *follows* from its purported set of premises. And in this case, because of its grotesqueness, our recognition was instantaneous, and thus funny.



Creationism Explained

Figure 2-1. Gary Larson's "Creationism Explained."

But obviously, and unfortunately, being able to pick up what grotesquely does not follow is not the same as being able to derive what subtly does follow. We might be naturally able to spot silly conclusions but we need to *learn* how to spot most of the interesting ones. [9] Or, to put it slightly differently, amassing the skills to assess the connection between premises and conclusions in most arguments – and especially in the sophisticated ones – will require work on your part. More precisely, in order to assess whether or not a conclusion follows, we need to understand, first, what *bonds* (so to speak) the latter to its given set of premises. And only after understanding the nature of this bond, we will be able to assess whether or not the it is a good one.

Luckily for us, we humans bond premises to conclusions correctly in only three ways. Or, to rephrase that, we know three ways to argue objectively in. We use *induction* (i.e. generalize) when we sample, for instance, thousands of dogs that happen to bark, and conclude that all dogs bark. We use *abduction* when we come up with the best explanation that can account for a given set of premises (like when we concluded, above, that there was a hungry beast on the other side of the door). And we *deduce* when we think in formulae, i.e., exploring the logical connections amongst the elements in the premises (like when we concluded that, if all dogs have some hair, and Fido is a dog, then Fido has some hair).

Now, there is a reason why it is important to be able to identify what kind of argument we are dealing with: each kind of argument has its own rules. What makes a good generalization (or inductive move) is certainly not the same set of criteria that makes, for instance, a good "best explanation" (or abductive move). In order to assess whether or not a given conclusion follows our premises, we will have to identify, first, what kind of argument we are dealing with:is it an induction, an abduction, or a deduction? Then we will have to apply the criteria of the identified argument "type" to the piece of reasoning in question. Once we have this under our belts, we will be able to spot factual claims; reconstruct the premises that led to them (which we learnt to do in the previous chapter); identify the kind of reasoning that was used (induction, deduction,or abduction); and assess it using the proper set of assessment criteria for that kind of argument. Only then will we finally have a fully operational fact-finding recipe! Let us start first with deductive moves.

---

[9] The fact that a great deal of North Americans (about half, to be precise, and according to most recent Gallop Polls) still cannot tell that Creationism is an implausible explanation sadly illustrates this point.

2.2 Deducing – or, How to Be a Good Vulcan

Maybe you know that in the Star Trek universe, Vulcans thrived on being exclusively *logical*. After watching a few shows, and if you felt like being technical and picky, you would realize that Vulcans, as good critical thinkers, don't restrict themselves to the use of *logic* and logic only, but also generalize, and come up with best explanations. To put it more precisely, *logic* is only a subset of critical; it is the name that we assign to the set of criteria that rules *deductive thinking*. To understanding this mode of critical thinking we will move next.

2.2.1 Formulaic Relationships

Deductions are infallible pieces of reasoning: once you grant the premises, the conclusion follows without a question. The unquestionable nature of deductions should not be mysterious. As we saw with the example of bachelors and mortals, in a very important sense, the conclusion was *already in the premises.* Mathematics, since it is a kind of deductive system, should help us see this point clearly. You should know that:

2 + 2 = 4

This is a *deductive* formula. It states that a set of 2's subjected to the rule of addition gives a 4. Or, to put it slightly differently: whenever you put two 2's together, you have a 4. See? The 4 *follows inevitably* because it was already (so to speak) *in* the two 2's. Let's go through it one more time, to be sure, but this time in the argument-reconstruction format we learned above. The mathematical formula above is a piece of reasoning (*deductive*, to precise); consequently, we can reconstruct it as an argument with premises and a conclusion in the following way:

P1: 2
P2: +
<u>P3: 2</u>
C: 4

It should be easy to see the elements and applied relations that this argument involves if we represent them with sticks and apply the relational rule (using the letter "I" for each stick):

P1: II
P2: addition (the kind of relational rule applied here)
P3: II
C: IIII

We know that the 4 inevitably follows from adding two 2's because our minds are picking up the good formula that is *underlying,* so to speak, the relationship between premises and conclusion. Let us now try to spot the same kind of backbone to our formula but this time in an argument made out of words instead of numbers. To do this, let us go back to one of the three reconstructed pieces of reasoning that we provided above. We argued that:

P1: If a given life form is mortal, then it must die
P2: Humans are mortal
C: Humans must die

Here, the premises state that death is some sort of unavoidable consequence of mortality. In other words, it claims that one cannot possibly have the one without the other. If you think about it, this argument is a no-brainer precisely because the relational backbone of its formula is almost on the surface. We are in a similar position as when we argue that, since Jean is a bachelorette, she cannot possibly be married. These moves are obviously unquestionable because there is a very short and infallible formula here: the "conclusion" is, in a sense, a restatement of the premises. If we know that somebody is mortal, then we already know (s)he will die, precisely because that is what mortal *means*. Along the same lines, once we know somebody is a bachelor, we also know that somebody is not married. The ideas of "mortal" and of "bachelor" *entail* the ideas of death and marriage, respectively. These examples might not be impressive, but what they should illustrate, again, is that once you have a good deduction, once you have a correct "formula," the conclusion *follows unavoidably*. But unavoidable formulae are not always so transparently displayed. Sometimes, even after a good reconstruction, the words we use obscure the relational move below.

When analyzing real arguments, it is easy to get distracted by the content of what we are saying and miss the abstract relationships that make the conclusion follow. To avoid this, and in order to make the elements and their relationships stand out, one can come up with symbols to stand for them. Let us try to reconstruct the argument about mortality using symbols for its elements and the relations amongst them. The first premise proposes a *class-- life form--* and a *property* that all members of this class have: *mortality*. The second premise asks us to identify a certain *sub-class: humans*. Naturally, your mind understands the abstract relations amongst the ideas of *class*, *subclass*, and their connection to the idea of *class properties*. Thanks to this understanding, your mind readily transfers the property of the class, in general, to all members of the subclass, in particular. So, let **M** stand for the class "Mortals;" **D** for the property "Death" belonging to all members of class M; and H for the subclass "Humans" that belong the class M. Translated into symbols, the argument above would read something like this:

P1: All members of class M have property D
P2: H is a subclass of M
C: H has property D

2.2.2 *Thinking in Relations*

What we are doing here is thinking deeply about what certain relationships mean, what they *entail*. And the idea of *entailment* is critical here. Pick a relationship; for instance, the relationship of *inclusion*. If you think about inclusion, you know that certain things must be the case whenever that relationship holds between any two things.

2.2.3 *Visualizing Relationships Using Venn Diagrams*

Venn Diagrams are very useful in order to visualize what is going on with the relational abstractions that go on below the surface in deductive moves such as the one above. These diagrams depict the relationships between *sets* of things. In them, sets are depicted as circles that enclose all the members belonging to that set. For example, the Venn diagram bellow illustrates that there are two sets, A and B, which have a third set of members in common, set C. This diagram, for instance, can be useful to illustrate the relationship of Canadian citizens (set A) and US citizens (set B) that share a common set of Canadian-US dual citizens (C).

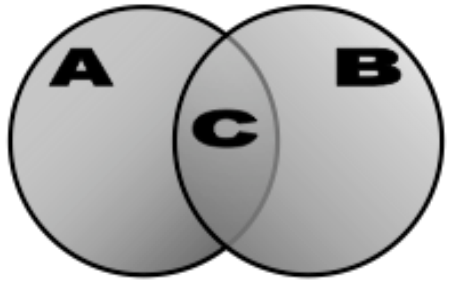


Figure 2-2. Venn Diagram illustrating that there are two sets, A and B, that have a third set, C, of members in common.

Our last reconstruction above could be depicted using Venn diagrams as follows:

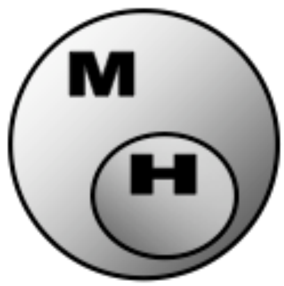


Figure 2-3. Venn diagram illustrating set M (for all mortal creatures) and its subset H (for humans). As the diagram depicts, since set H is a subset of set M, the property D in question (death) belongs to H as well as to every other member of M.

We have thus used Venn diagrams to understand the abstract relationships between the elements in the premises and the conclusion. Now we will use them to generalize a bit and try to understand deductive moves in general, that is, to understand the relationship between the premises (not just its elements) and the conclusion, in *all* deductive moves.

2.2.4 Venn Diagrams, the Valid, and the Logical

We said before that in deductive reasoning, the conclusion re-states the premises. A better way of putting this would be that the conclusion was already harbored in the elements, and the relations between those elements, in the premises themselves. Or, to put it in yet another way, all you need to do to extract deductive conclusions from a given set of premises is to “release” one of the many good formulae that the symbols and relationships deployed in the premises allow. Back to math, for the simplest examples, we can say that (with the elements and relations italicized) one can deduce a *6* out of the *addition* of two 3’s as well as out of the *addition* of a *2* and a *4*. Similarly, we have also argued that:

P1: All mammals have some hair
P2: Dogs are mammals
P3: Fido is a dog
C: Fido has some hair

In a Venn Diagram:

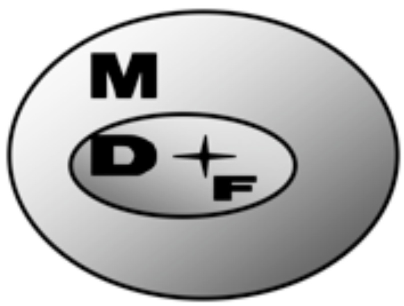


Figure 2-4. Venn diagram illustrating set M (for all mammals) and its subset D (for dogs). As the diagram depicts, since set D is a subset of set M, and F (for Fido) is a member of the latter, the property SH (for some hair) for the set M belongs to F as well as to every other member of M as well as of D.

At this point you might be able to see that in all of the Venn diagrams above, deductive conclusions are always *within* the set of their premises. This relationship is what I also tried to capture above by using words such as *entail*, *within*, or *inside*. The point here is that whenever you deduce a conclusion from a given set of premises, you never stray outside of what the premises grant, or entail. This pattern, in a generic diagram for deductive moves in general, should look something like this:

Figure 2-5. Venn diagram illustrating the case of every good deductive move, where set P is all elements and relationships in a deductive move's set of premises and C is its subset (one of its possible conclusions).

Now, when this relationship holds--when a conclusion, in a deductive move, keeps within the elements and relationships that are established in the premises-- we say that the move was *valid*.[10] In the Venn diagram above, we illustrated only one of the valid deductive conclusions (C1) that one could possibly make while still respecting these two conditions (that is, sticking to the premises' established relational rules as well as its elements). But again, in principle, the number of *good formulae* that one could possibly create --that is, those that respect those two elements-- is unthinkably large.

It is also time to introduce another technical term that will prove very helpful: *logic* (the term that, as I mentioned above, was misused by Star Trek's Vulcans). The sentence above introduced, in a rather sly way --sorry!-- the concept of a *good formula*. Again, and a bit more precisely, a good formula is one that falls within the permissible moves of a given deductive system. (Remember that, in math, for instance, good formulae were those that respected the established symbols, as well as the established rules of permutation for the symbols (e.g., addition, subtraction etc.) *Logic, then,* is the sum total of these good formulae. Of course, the sum total of these good formulae does not exist anywhere, definitely not in somebody's mind. This infinite set is a subject of exploration and study, the subject study of Logicians, to be more precise. So, putting it all together, Logicians study

[10] It is important to bear in mind that, from now on, every time we use the word *valid*, we are using it in this very precise, technical meaning – and not in its colloquial one (to denote a point that is correct, or well taken).

valid deductive moves: most specifically, those that stick to *logic,* or the set of allowed relational rules and symbols.

2.2.5 Pumping Facts

The points we have just made might look mighty abstract (and technically speaking, they are, of course!). But if you think that this makes them irrelevant to your life, you could not be more mistaken. As I mentioned before, the valid deductive moves that one can make from a given set of premises is vast: think, for instance, about all of the possible good formulae that one could derive from the set of real numbers and their allowed relations. Now, what would happen if we bring the idea of content back into this picture and ask the following question:

> What would happen if one were to derive all possible *valid* conclusions from a set of *factual* premises?

First, if your deductions are valid *and also* come from factual premises, then we call their conclusions *sound*, that is, valid *plus* factual). Of course, you could not possibly have enough time to derive *all* possible valid conclusions, but if I were to give you a lot of time, and if you were very dedicated, imagine all the deductive conclusions that you could pump out of such a fruitful system! Now remember that a valid deduction stays well within the realm allowed by its premises – if it were ever to stray outside of it, then the conclusion would be *invalid*. Then (and the practical import of this cannot be exaggerated): *if you start from a set of factual premises, then **every valid conclusion you derive must also be factual (that is. sound)**.* So to put it very succinctly, if you start with facts, valid deductive moves will only pump out more facts, nothing else. And you can bet that some of them, at least, will be very surprising. Let me illustrate this with a real example that changed the world.

2.2.6 Darwin Deduces the Fact of Natural Selection

In 1859, Charles Darwin published *The Origin of Species*, a book that almost single-handedly created the discipline of the biological sciences and changed our understanding of (us) humans, in particular, and of the nature of life, in general.[11] At the core of his book was Darwin's discovery of

---

[11] For everything on Darwin (including all of his publications) you should go to the UK's fabulous website: http://darwin-online.org.uk/.

natural selection: the natural mechanism that shaped (without the intervention of a some mind – divine or otherwise) the wonderful adaptations that life forms present and that had so marveled naturalists for ages. This discovery was at least as momentous as Sir Isaac Newton's discovery and understanding of gravitational forces about a hundred years earlier. Just like Sir Newton with the force of gravity, Darwin *deduced* the inevitability of natural selection from a set of unquestionable facts about the way our world is. To put it in more technical terms, Darwin concluded natural selection out of a *sound deduction*.

Now that you understand the way sound deductive moves work, you should be able to grasp the significance of Darwin's feat. As we have seen above, in a sound deduction, the conclusion follows inevitably from the premises and, in a sense, was already *contained* within them. Therefore, the factual status of the latter must necessarily be transferred to the former. In other words, once you grant the factual status of your premises (which any sensible person must, because they are –-in this case-- unquestionable), then you *must* also grant the factual status of the conclusion. we have also seen that how much one wishes to believe or not believe in the conclusion does not really matter. Darwin's deductive move demonstrates, as all sound deductive moves do, that natural selection follows *inevitably* from the way the world is. And, unfortunately, it does this whether we like it, or believe in it, or not. In this sense, precisely because of their disregard for the feelings of the subject doing the thinking, deductive moves are objective -- and thus not very user friendly. But since we have also seen that not getting one's facts right can easily kill you, it is good practice to teach oneself to do this kind of good thinking naturally, and to pay close attention to good thinking as you do it. In other words, one should learn to *appreciate* the unfriendly nature of deductive moves! So, let us now jump into Darwin's deduction of natural selection itself to cement our understanding of the nature and force of this type of reasoning.

2.2.6.1 The Way the World Is: Getting the Factual Premises Right

Darwin first notes that there are a number of unquestionable facts about the world. Mind you that this unquestionability should not be taken to mean that one can spot these facts from a favorite armchair. Darwin's factual premises are easily accessible and uncontroversial to anybody that should care to do the work necessary to adopt an objective stance and look at them [straight on]. This should become evident once you see the factual premises that Darwin considered.

2.2.6.1.1 Adding Limited Resources

This one is easy. Everybody knows that there is not enough food, shelter, mates, or, in short, of all that a [everything any] creature would require in order to produce babies. These bare necessities, taken together, are technically called *resources*. This is sometimes called the Malthusian principle, named for Thomas Robert Malthus, the British politics and economics scholar of the late 1700s, who noted that limited resources was an unavoidable condition affecting every human population. As soon as Darwin read Malthus' ***An Essay on the Principle of Population*** (at the time, a very popular read), he immediately saw the broader necessity of this principle; this realization precipitated the final discovery of natural selection. Here, we will call this principle the Principle of Limited Resources.

2.2.6.1.2 Superfecundity

The second factual premise is technically called *superfecundity*, and put rather simply; it means that there is an excess of creatures being born. Or, to put it in yet another way, not all creatures that are born can possibly survive. One can easily see this with flies, for example, but Darwin did not take any chances. He provided a huge number of examples among various species and also picked up the slowest breeder known in the planet, elephants, to calculate that, if all elephant babies produced were to survive, it would take 500 years to cover the entire planet with elephants. Since no herd of surplus elephants is squashing you right now, nor are there any extra flies, mice, dogs, or sparrows (to mention just a very few possible species) choking the planet, it is obviously the case that not every baby produced survives. Let us put these two factual premises together and deduce our first unavoidable conclusion.

P1: Given limited resources
(when the amount of bare necessities is lower than that required by the number of babies around)
P2: Given superfecundity
(when the number of babies around is greater than that for which the available bare necessities would suffice)
C1: There must be a *struggle for existence*

By *struggle for existence,* Darwin meant, rather simply, that life for each one of these babies was not going to be easy. As each one of these new creatures strove to become an adult and reproduce itself, it would have to compete (directly or indirectly) for the limited resources available.

2.2.6.1.3 Heritable Variability

Variation is a fact of life. Anybody that cares to take a look should notice that, for every batch of babies, it is almost certain that none of them will turn out to be identical. And mind you, this is the case even for identical twins, that in truth are only *almost* (genetically) identical. Noweverybody also knows that some of this variation will turn out to be heritable. The science behind heritability was, in Darwin's time, non-existent. Gregor Johann Mendel, the discoverer of the aptly named laws of Mendelian Genetics, was a contemporary of Darwin. But unfortunately for both these great scientists, Mendel's findings only became widespread almost a century *after* he discovered them. I will remind you of this case later on, when we discuss the massive role that scientific societies (with their codes, journals, and meetings) play in the development and progress of scientific knowledge.

So even though in Darwin's time nobody really *understood* what was going on behind heritable variation, neither would anyone in their right mind dispute its factual nature. Furthermore, Darwin was able to establish, through methodical and exhaustive inquiries to the genetic engineers of his time (breeders of domestic varieties of plants and animals), that heritable variation was readily available; that it was small in nature (you almost certainly could not get wings on a pig, no matter how long you waited); and that it was, most unfortunately, random. Imagine how convenient it would have been for, let us say, the breeder of dog varieties if they could direct the nature of heritable variation, to cut the tails of those Boxer puppies and see how, in the next generation, they start coming up with the desired knobby tail. But alas, everybody involved in the breeding of domestic varieties knew that that was not the case. What they knew for a fact though was that, amongst all the variation available in babies, there was always a rather large pool of random, small, heritable variation.

2.2.6.1.4 Natural Selection

Now, if one uses the previous conclusion as a premise and adds the fact of the nature of heritable variation, the following can be concluded.

> P1: Given the *struggle for existence*
>
> P2: Given the availability of numerous, small, random heritable variation
>
> C2: There will very, very occasionally be a mall piece of random variation that gives its possessor an edge in the struggle, and is also heritable

Let us call this very special and rare kind of heritable variation *adaptive variation*, and let us combine it with the (obviously, also unquestionable) fact of *time*.

> P1: Given adaptive variation
>
> P2: Given time
>
> C3: Adaptive variation will tend to accumulate creatures within a given lineage, furnishing them with traits that are helpful in the struggle for existence

So, what Darwin discovered thus far in this process was that, in a universe where there are superfecundity, limited resources, heritable variation, and time, creatures will tend to get designed to be good strugglers --whatever the nature of that struggle happens to be in their lineage. Darwin called this principle, or force of nature, *natural selection*.

It would be hard to exaggerate the historical significance of this deduction. The amazing discovery here was that there is a force of nature out there that could produce amazingly intricate design but without a mind, a purpose, or any foresight whatsoever. This, as I am sure you can appreciate, is a monumentally surprising discovery. In our human daily experience, design requires a designer. And the better the design, the better, more intelligent the designer must be. Creatures are so well-designed that, still today, some people cannot relinquish their intuitive, everyday understanding of the designee-designer process and so insist that the amazing design of living creatures requires the existence of an all-powerful and intelligent designer. But you should now be able to understand that the force of deduction completely takes the air out of their insistence. In a universe where the properties outlined above (the premises used in Darwin's long argument) are a fact, natural selection--unavoidably, unquestionably, and without a doubt-- is also a fact: even though we had to wait for Darwin to figure and spell it out, embedded as it was in the list of factual, unquestionable premises above, natural selection was always there, waiting for Darwin while it stared us in the face.

This is, of course, the method that the Glieseans identified as *the Pure-Reason Method*. I hope this dramatic example illustrates the force with which the factual nature of a premise can be automatically transferred to a valid conclusion. Thus, deductions can produce previously-unknown and surprising new facts from obvious boring ones.

But there is another common use for deductive moves, one that is *so* common and *so* salient within our human experience that, along with our unlucky Gliesean descendants, we were mislead to categorize it as an entirely different method: *the Trial and Error Method.*

2.2.7 Testing: Deduction and Trial and Error

2.2.7.1 The Certainty of Refutation

If one thinks about trial and error carefully, one should realize that it boils down to a deductive move: whenever we perform a trial, or *test*, we are simply arguing that a certain result (by logical derivation from the premises) *necessarily means* a certain conclusion. Thus, in a valid deductive test, the premises entail the conclusion. Let us try to see this with the help of our most unfortunate example from above.

> P1: If Gliesean Tarantulas are good pets, then they should be nice and cuddly when we try to pet them.
> P2: They are *not* nice and cuddly when we try to pet them.
> C: Gliesean Tarantulas are not good pets

The conclusive fact that these alien creatures are not good pets was derived from our mental formula that worked out what good-pethooded-ness (to call it somehow) necessarily entails. As usual, let us get rid of all the concreteness in this piece of reasoning so we can unveil its formulaic nature. We will convert "***Gliesean Tarantulas are good pets***" to "Δ;"

“***then***” to “$\Rightarrow$;” “***they should be nice and cuddly when we try to pet them***” to “α;” and “***no***” to “~.”

$$\begin{array}{ll} \text{P1: If } \Delta \Rightarrow \alpha & \\ \underline{\text{P2:} \quad \sim\alpha \quad} & \\ \text{C:} \quad \sim \Delta & \end{array}$$

I hope you can see that this formula is valid, as evidenced, for instance, by the fact that once you understand what the symbols are saying, you agree because you *cannot* think otherwise. Just to make sure, let us add a bit of a narrative to express what this formula is saying: whenever there is an Δ then necessarily, unavoidably, there must also be an α; if it just so happens that α is nowhere to be found, then there could not possibly be an Δ around either. This move should be obvious and not surprising, but its implications are mindboggling.

We have hit upon a major, monumental discovery. By trying to understand what was going on when we apply the deductive method of fact-finding, we discovered our first piece of certainty. If our premises are true (a *huge* caveat) then a failed test gives you certainty that something went (deductively) wrong. Again, we should now understand that when a test fails we have a deductive proof of that failure. To put it slightly differently, we can prove that we were wrong. Technically, we call this kind of proof a *refutation*. Now can we do the same when a test succeeds? Or putting slightly differently, when a test pans out, can we thus prove that we were right? Sorry, but no (such luck). Let us understand why.

2.2.7.2 The Support and Justified Hope in Confirmation

Let us momentarily postpone the tragic loss of our fact-finding committee member. Imagine that the Gliesean Tarantula was nice. When our adventurous but foolish descendent went to hug the creature, the creature made a purring sound and gently “embraced” our committee member in return (or whatever you would call a hug-like motion executed with spider-like mandibular appendages). This time the formula, which in the trade we call a *confirmation*, would read as follows.

$$
\begin{array}{l}
\text{P1: If } \Delta \Rightarrow \alpha \\
\underline{\text{P2:} \quad \alpha \qquad} \ \textbf{(?)} \\
\text{C:} \quad \Delta
\end{array}
$$

Is this a valid move? Many of us are readily moved by positive test results and would say *yes*. We primates are hopeful testers. Probably, those ancestors of ours that were not, on average, were less successful at leaving progeny than the hopeful ones. In any event, it is interesting to note that most people that see this formula take it at face value as valid. But, on careful consideration, reason wins and they see the problem. Do you see it?

It might be easier if we add a narrative. This formula claims that whenever there is an Δ then necessarily, unavoidably, there must also be an α. So far so good, but then it adds that, this being the case, if α is found, then one must conclude necessarily that there is an Δ around. This, I hope you can see, is wrong: it would be flipping the meaning of the first premise around! The first premise tells us that Δ is unavoidably followed by α, not that α unavoidably followed Δ. Certainly, if you are a reasonable person, your hopes would go up upon encountering α. But even though justified hopes should be taken very seriously, they are far from being a *proof*. The first premise (that $\Delta \Rightarrow \alpha$) does not say anything about what else might imply α. Back to our example above, there are countless possible alternatives that could have been responsible for the results above. Gliesean Tarantulas could do their "hug and purr" thing as a prelude for attacking, dying, falling asleep, calling for reinforcement, entering a feeding frenzy, hypnotizing their prey, or mating. Consequently, the premises did not entail our conclusion and this was thus an *invalid* move. To put it slightly differently, there is no certainty in confirmation; just the justified hope that comes from 1) knowing that there was no refutation (where certainty lies) and 2) the possible gathering of additional support that other pieces of reasoning can provide *in addition* to the test in question.

So by studying deductive moves we now understand that we can certainly disprove facts, but only confirm (not prove) them. Some of you must have heard something similar when people talk about science. "Science can never prove anything -- only disprove hypotheses and theories" -- they say. When you heard this, you might have felt pity for those poor scientists. Your pity should now extend to humanity in general (yourself included, of course). But more, much more about this when we get to our section on understanding science. For now, let us continue assessing the fact-finding recipes.

As we discovered above, deduction (*the Pure-Reason* and *Trial & Error Methods*) is not the only way we know to objectively tie premises to conclusions: there is also *induction* (*the Category Method*, when we generalize from a given set of premises) and *abduction*, (*the Best-Explanation Method*, when we come up with the best explanation that can account for a given set of premises). We will move next to these two, each in turn.

3 CHAPTER 3 -- CONCLUDING FACTS VIA INDUCTION

Inducing and abducing conclusions from premises have an important aspect in common that they do not share with deductive moves. We saw that deductive moves were certain: if the move is valid, and the premises are true, then the conclusion *must also* be true. As we argued above, this is because, when one deduces, one is exploring what the premises *entail*, that is, to put it slightly differently, what is already *within* them.

However, when we generalize, or abduce, we are leaving the safe, established realm of the premises and adopting our most objective stance possible, in order to come up with what the premises most probably *indicate* (to put it somehow). In this sense, the force tying premises to conclusion, in both abductive and inductive moves, is not as strong as in the case of deductions. To put it in yet another way, one could have a set of factual, well-established premises that are true, produced from them an excellent inductive or abductive move, and still find out later that these conclusions are false. Putting it shortly, the truth-value of the premises, both in an induction and in an abduction, *do not guarantee* the truth value of the conclusions. This should be easy to understand with the help of a Venn diagram.

Figure 3-1. The Venn diagram on the left illustrates the relationship between premises and conclusion in a deductive move. The diagram on the right demonstrates the same for either an abductive or an inductive move. It should be evident that in the latter case, the set containing the premises can be true while the conclusion is not, since the conclusion lies *outside* of the set containing the premises. The same cannot possibly be the case in a deductive move (where the set of the premises *contains* the conclusion).

When we abduce and induce we are making a leap, but not a leap of faith. It is a leap that should be guided by the fulfillment of the criteria that make an excellent inductive or abductive move. So what are these criteria that make an excellent inductive or abductive move? To these we will move now, beginning with inductive moves.

3.1 Inducing from Particulars to Universals

You must have experienced the following at least a couple of times. You are thinking about a friend, and at that very moment, that very same friend calls you on the phone. Most of us can not help the temptation of considering this to be more than a mere coincidence, something that happens an isolated number of times, for no good rhyme or reason. In fact, we jump to the opposite conclusion and consider that there is a happening here. But do we? In order to assess this problem, we will need some useful terms.

Very broadly, we will say that when we are dealing with a set that, let us say, has all members of the kind *dog*; we are dealing with the *universal* set “dogs.” This set then, by definition, has all the dogs ever (past, present, and future). In this set, each individual dog will be thought of as a *particular* instance, or member of the universal set. In our example before, Fido is a particular member of the universal set “dogs.” Now we can see that, when we *generalize*, we are moving form a number of particulars to propose (or *conclude*) a given universal. Usually, what we are also doing is talking about some kind of property that has more than one possible state of being, or value. Think, for instance, about the property *barking*, some dogs bark a lot, other little, and a few not at all. To be a bit more technical and proper, we should say that these are all different possible values for the *variable* barking. Some variables have only two possible states or value, like *some-hair* versus *no-hair*. An excellent induction would be that all dogs (like any other mammal) have some hair (shaving your dog completely is cheating).

There are many kinds of useful universals. The universal set that houses all the particular instances of some happening or other we usually call a *pattern*. For instance, they say that the people of Konigsberg used to set their watches following the famous philosopher Immanuel Kant's (1724-1804) regular afternoon walks. So, we can say that Kant's afternoon walks were highly *patterned*. Or, to put it slightly differently, we can say that there is a universal set that holds all the particular afternoon walks of Kant as its instances. When the pattern, or regularity happens in nature, we will call it a *regularity of nature;* or *phenomenon* (*phenomena* in plural form). We know of many *phenomena* out there: like, for instance, the tides; the seasons; the way life is broken down (most of the time) into species; or the climate (at least as it used to be for millennia, before global climate change set in).

We can now ask the question above a bit more precisely. When your friend calls exactly at the same time you were thinking of her, should we consider this to be a particular instance in some universal pattern (in this case, some mysterious phenomenon of clairvoyance, or telepathy)? Let us reconstruct the argument we would be making for this remarkable conclusion.

P1: Telepathy-like event 1 (of 12/3/04)
P2: Telepathy-like event 2 (of 1/6/05)
P3: Telepathy-like event 3 (of 10/3/06)
P4: Telepathy-like event 4 (of 2/23/08)
P5: Telepathy-like event 5 (of 1/30/08)
P6: Telepathy-like event 6 (of 6/3/09)
C1: Telepathic-like communication is a phenomenon

You might be (a bit?) moved by this set of premises to accept the conclusion, but (hopefully) you are not convinced entirely. It is our good fortune that we are all rudimentarily skilled at spotting what makes a good argument for patterns, and what doesn't. The most obvious criterion for excellence here deals with brute quantity: everybody knows that one particular hardly amounts to good grounds for concluding a universal. This seems to be one of those cases where the more the merrier: the more particulars accumulated (this is what scientists call *samples*), the stronger our generalization seems to be. We can call this an induction's *sample size*.

But how good are six particulars (like in the example above) for a sample size?

This is a very difficult question. Basically, what we are asking here is: is our sample size close enough to the universal? I hope you realize that that is precisely what, form the start, we do not know! Furthermore, when we are dealing with a true universal statement (like, for instance, all dogs ever) our sample will never be close enough (how close can one get to *infinite*?). So, when we make an induction we bite this bullet and use some (objective) statistical maneuver that gives us an indication of how close we should trust to be. Explaining these statistical maneuvers is, of course, way beyond the scope of this essay. So, for our present treatment, it should suffice if you understand the relationship between sample size and the universal that it is trying to capture.

Now, it is important to realize that sample size is not the only criterion that determines whether or not we are dealing with a good inductive move. Quantity, as we all well know, is not all that matters. A simple example should help us drive this point home. Imagine I tell you that I have sampled 80,000 Americans and asked them if they have ever tasted moose meat, and that I found out that 64,000 of them report they did (imagine also that you are kind enough to believe me). Now, if sample size was all that matters, one could conclude inductively that about 75% (the proportion of affirmative answers in my sample's total) of all *Americans* have tasted moose meat. Now, if you are reading this in, let us say, the states of Florida, Texas, or California, you might be rather surprised. Since it would most certainly be the case that, in your experience, it is not the case that an average of 7.5 out of each 10 persons you meet has tasted moose meat. However, if you are reading this in the state of Alaska, were my fictional (though probably quite appropriate) sampling took place, you would find this inductive conclusion very true. The lesson to learn here is that sample size matters greatly, but *representativeness* matters greatly also. In other words, in a generalization, we want to make sure that the samples we are picking are not restricted to some quirky area of the universal we are trying to grasp. Ideally, we want our sample to give us a good picture of all the nooks and crannies in the universal. That is why, since one does not know where the quirky nooks and crannies lie, that we *randomize* our sample collection. Or to put it more simply, we collect our particulars trying to avoid any personal biases that one might have.

For instance, to continue with our fictional example above, even though I used the sample size that statisticians recommended, and since I was going through a particular lazy stretch of the month, I decided to pick all my samples close to home (in Fairbanks, Alaska). That is why my conclusion about *all Americans* was so biased. We should now understand that, in spite of the fact that my induction had a good sample size (that which statisticians recommended given my intended universal), my sample was not *representative*, or to put it slightly differently, it was *biased*.

So, in conclusion, when we induce, the criteria of goodness that we keep an eye on are *both* sample size *and* representativeness. Now, before we move into our next type of argument (abductive), we need to think a bit more about what is it exactly that we are concluding when we make an inductive move.

3.2 Correlation

A recent study[12] found that men that shaved less than daily, when compared with men that did, have a 52 percent higher chance of dying of cardiovascular disease. Please do not throw that shaving kit out the window just yet! If you already did, I am sorry, but it means that you misunderstood what inductive moves (like this one) have to tell us. This study followed 2,438 men for 20 years. Statisticians and MD's would both assure you that the sample size is adequately large and not biased. This is a good induction and is telling us that the variable shaving *inversely co-varies*, or is *inversely correlated* to the variable of dying of cardiovascular disease. The idea of *correlation* simply points out to the fact that the variables in question seem to change in tandem. Alcohol and drunkenness go hand in hand; co-vary together; the more alcohol one drinks, the more drunk one gets. So we can say that alcohol and drunkenness are correlated. Another correlation with the amount of alcohol drunk is the lack of coordination. In this case, as one variable goes up (alcohol) the other (coordination) goes down; so we say that the variables are *inversely* correlated.

Naturally, when noticing inductive patterns like these, your brain always goes to go further. It is never enough to discover that two variables co-vary in some way or other. We all want to jump to the idea that one of the variables is *causing* the other to vary one way or the other. The word

[12] Ebrahim S, Davey Smith G, May M, et al. Shaving, coronary heart disease, and stroke: The Caerphilly Study. Am J Epidemiol 2003;157:234–8.

correlation was coined to alert us of this very problem. Noticing a regularity or pattern says almost nothing about *causation*. Naturally, if A *causes* B, then A and B will go hand in hand, and we might be lucky enough to induce this pattern via the recording of many instances of A's followed by B's. It is true that wherever there is a causal relation between variables there is co-variance between them. But not everything that goes hand in hand, or co-varies together does so because one is the cause of the other. Or, to put it more simply; co-variance does not necessarily mean causation. Not to note this, is a very common, and dangerous mistake.

Sometimes, co-variance is just a coincidence – something that a good induction will spot out for you. For instance, good inductive moves demonstrate that our previous telepathy-like putative phenomenon should not be taken seriously. Statistics alone (that is; sophisticated inductive moves) will predict that such telepathy-like events are likely to happen. In their book, *Debunked!* (Johns Hopkins University Press, 2004), Georges Charpak and Henri Broch argue, for instance, that, in the United States alone, with a population of 295 million, events with one-in-a-million odds should occur about 295 times a day! Given the number of friends you have, the windows of opportunity for you to think about them, and for them to think about you, make the chances of some of them actually calling you while you are both thinking of each other not at all miraculous.

Some other times, the simple fact of co-variance only indicates that a third variable, the actual cause of the whole pattern, is working in the background, unnoticed. Or, to put it slightly differently, in these cases, the reason why two given variables are correlated is because they have a cause in common. Something like this is the case with our correlation above between shaving and coronary death. As you might have suspected, it is highly implausible that shaving can cause coronary disease. What scientists studying this phenomenon suspect, is that sex hormones are responsible most of the discovered patterns. It just so happens that the men that shave less than daily and show a high correlate with coronary death, also show the correlates of having breasts that are tender and woman-like; low facial hair growth; low sex drive; and they also tend to be single; and to smoke.[13]

[13] I hope you picked this one up and wondered if smoke alone wouldn't explain coronary death. It could, so the researchers used statistical devices that crunched

3.3 Certainty with Induction?

The short answer is *no*. Inductions can be weak, or strong, they can even be superbly strong, but unlike deductions, they do not pump the truth-value of their premises to their conclusion. An excellent induction, which sample is truthful and unbiased, including millions of black ravens, will lead to an excellently strong conclusion that All ravens are black. But no matter how large and unbiased its sample could be, the truth of the premises will never *guaranty* the truth-value of its conclusion. This should be easier to see with the help of the Venn diagram below. Since we are leaping from subset to inclusive set, the truth of the first cannot possibly guaranty the truth of the second.

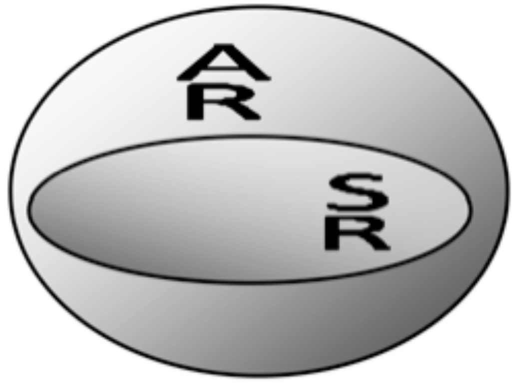


Figure 3-2. Venn diagram illustrating the set of all ravens including, as a subset, the unbiased, excellently large set of premises made up of all the ravens sampled.

But, you might *ask*, what if one *tests* the conclusion that *All ravens are black* via a deductive move; would not that guaranty the truth-value of the premises? More precisely, you argue the following.

P1: if All ravens are black ⇒ next raven will be black
P2: Next raven is black ______________________ (?)
C: All ravens are black

Or:

their data in such a way that it tested whether or not this could be the case. It was not.

P1: if All ravens are black $\Rightarrow$ next raven will be black
<u>P2: Next raven is white</u>
C: ~ All ravens are black

It should be easy to see now that this would not help much and that we still have no certainty here. If the test is confirmed (the next raven you see is black), then, as we learnt in the previous chapter, we can trust, but the move is not valid so we have no proof (thus the question mark over the conclusion line). If the test is negative (the next raven you see is not black), then you prove the inductive conclusion false. So, the best we can do, combining excellent inductions and excellent tests are superbly confirmed correlations or certain refutations. So, we will never have proven correlations, but our excellently supported ones should do.

Still, our most excellently supported correlations cannot give us the causes lying behind them. Inductions are mute regarding causes. Deductions can surely test possible suspects once we have them. But we need to explore our third and last kind of reasoning in order to understand how to *produce* them. Obviously, finding responsible causes out is at least as vital as finding the patterns or correlations in the first place. Without knowing the causes behind the patterns disclosed via induction, and tested via deduction, without understanding what is going on behind these phenomena, we do not stand a chance to be able to deal with them. No amount of inductive moves will help us do this. For finding out causes, we need the help of our third and last kind of reasoning: *The Best Explanation Method,* or *abduction*.

4 CHAPTER 4 -- ABDUCTION: CONCLUDING FACTS VIA INFERNCES TO THE BEST EXPLANATION

If the Vulcan Mr. Spock was the (confused) pop-culture icon for deductive moves and their logic, Sherlock Holmes should be the most appropriate (albeit not confused) one for abduction. In my experience, the word "abduction" brings about thoughts of little green men with buggy eyes and unwholesome intentions. So I hope to dispel these distracting associations by changing the name for this kind of reasoning to a more descriptive one. From now on, we will call abductions *inferences to the best explanation.* But, since this is name is way too long, we will shorten it to *ITBE* (which of course stands for the initials in the proposed new name).

Now, before we proceed any further, we need to think a bit about a key term in our new name: *explanation*. Our daily meaning of this term is quite adequate, but it packs much more than what we usually notice. So let us start unpacking it right now.

4.1 Describing versus Explaining

Our species is so fond of inferring to the best explanation that we can barely hold our horses when we are confronted with an opportunity to do it. Descriptions, albeit of course important, do not seem to be met with such enthusiasm. Nonetheless, they do tend to work as some sort of irresistible prelude, or invitation for the first. When we *describe* we provide statements about the way the world *is*. But when we *explain* we do something completely different: we provide a *reason why*; we provide the *cause* of whatever way we described the world to be in. Descriptions and explanations are so different that they even have their own dedicated words to elicit them. When we request a description we ask *what* but in the case of an explanation we ask *why*. Let me illustrate this with an example.

Let us say that Jane has an appointment with me at 10:00 am, but that she shows up late, at 10:30 am. Since I am a bit frustrated by this, I request an explanation; I ask her *why* is she late. Jane goes on to tell me that she is late because our appointment was at 10:00 am, but she arrived half an hour later, at 10:30 am. When I insist that she has not yet *explained* to me *why* she is late, Jane proceeds to tell me (now a bit frustrated herself) that she is late *because* our appointment was supposed to happen when the long hand of the clock was at the number twelve and the short one at the number ten, but that she actually arrived when the long one was at the number 6.

As you can see, the problem here is that Jane is misusing the word "because" in her sentence. Jane is disregarding the difference between describing and explaining that we are trying to establish here. When one *explains*, when one answers a *why* question, one is not providing an account of some state of affairs or other, one is coming with the cause of the happening. The fact that, colloquially, most explanations start with the word *because* should be very telling!

We should are now ready to see how inductions and ITBE's complement each other. We saw, in the previous chapter, how inductions help us come up with a particularly difficult kind of descriptions out there: those of patterns. But we also saw how induction was helpless when trying to discover who has done it; or which was the variable responsible (that is the cause) in setups where more than one variable were co varying together. When the problem is explaining, not describing, we resort to ITBE's.

4.2 Cause-Finding

4.2.1 Red Rain

On July 25th, 2001, something strange happened in the state of Kerala, in India. Very early that morning though, a very loud thunder, accompanied by a flash of lightning, was heard. July is within the monsoon season, but thunderstorms are not at all usual during this part of it. So this was unusual, but what happened next was even more. The lightning and thunder were followed by thick, downpour of *red* rain.

I bet you that in the process of reading this short paragraph you already went through at least two different ITBE's. There are two interesting things I hope you can notice in this automatic thinking that you performed. First, how eager is your mind to do it -- that is, to come up with explanations for descriptions of events. And second, how methodical your mind is when it sets off in this direction. It is very likely that, if you (sort to speak) *rewind* the tape of your thinking, you would see that, first, and quite naturally, you concluded that the best explanation for the first string of events (before you heard of the red rain) was a thunderstorm. More precisely:

1) You noted the events (premises) that needed explanation -- that is: a loud bang in the sky accompanied by a flash of lightning.

2) You added a silent, albeit huge package of premises that provide the background, relevant knowledge to event premises above -- that is, for instance, all the premises regarding your rudimentary knowledge about climatology.

3) You concluded that this compounded set of premises (events at hand plus background knowledge) was *best explained* by a thunderstorm. Or, to put it slightly differently, your cause-finding thinking homed in on *thunderstorm* as the most plausible *cause* for this particular set.

Now, above, I said that you went through at least two best explanations as you read the short first paragraph of this section and I have just described the first. You read the premises of lightning and thunder, and you concluded thunderstorm. But everything changed, I mean, our best explanation when out the window, discarded as implausible, when I mentioned the *red rain*. Let us try to understand what happened in more detail -- it will give us useful insights into the nature of this type of inferences and, more importantly, into what makes a *good* ITBE.

The *thunderstorm* best explanation was discarded when a new premise was added, the red rain premise. If the new premise had been plain old rain (that is rain that was not of a red color), there would have been no need to drop our best explanation because, in a sense, it (the good old rain) would have been *allowed* by it. Or, to put it a bit more precisely, the background climatology knowledge package that we added to the particular events set (thunder and lightning) would have specified that the added event, rain would be yet another expected effect of the cause we had picked (thunderstorm). But the background knowledge package does not specify *red* rain! If you are like most of us, the addition of *red* would have been so much *not* specified by the background knowledge set, that your mind would have (rather desperately) gone out on a limb and start proposing wacky explanations, such as aliens, or divine intervention.

Please stop. Let us rewind again and try to now notice and investigate what is it exactly that determines that aliens (or divine intervention) are far-fetched, or outright wacky explanations? Or, to put it more precisely, what are the criteria that one uses to determine what makes an explanation the poorest, or the best one?

4.2.2 The Power of Background Knowledge

For the sake of our discussion, it would be useful at this point to think of ITBE moves as if having two kinds of premises. The first kind deals with the phenomena that we are trying to explain and its allied events. Let us call this kind: *phenomena* premises. The second kind we have already identified as the background knowledge package. We also have already noticed that it was the background knowledge package that sent us looking for a new explanation when a new phenomena premise (the red-rain premise, in our discussion) was added to our argument. Let us now try to understand how does the background knowledge package play this crucial role in determining what is acceptable, or not, when it comes to picking a would-be explanation.

Unlike the phenomena premises, that (sort to speak) provide us with the crime scene, the background knowledge package gives us all the relevant forensic procedures and science, and a list of the usual suspects. In less metaphorical terms, this package contains all the relevant conclusions that had been reached by other trusted inductions, deductions, and inferences to the best explanations regarding the particular set of phenomena premises at hand.

Going back to our example, as we read the list of phenomena premises, it was as if our mind, quickly and inadvertently, did two things: one, *categorize* the kind of phenomena we seemed to be dealing with (a climatology one); and two, retrieved and *apply* all relevant information in this category to help us determine what *made more sense* (sort to speak) as the best explanation. When we got to the red rain, the same package sent (again, sort to speak) a red alert. Even our humdrum background climatological package (the one with all the every-day good inductions, deductions, and abductions about weather phenomena that we all have) tells is an oddity that the package cannot deal with. Messing up either of these two-steps (categorizing and application of background knowledge) will make our ITBE's go awry. As if to demonstrate that real life is always better than fiction, the real-life efforts to explain the red-rain phenomenon drives this point.

As you probably have already guessed, red rain is not a common part of any well-known climatological phenomena. So, the first step, that of categorizing the phenomena premises set, is not a straightforward one. A mistake made at this stage would be crucial in the sense that we would not be quite sure about which background information one should bring in, in order to judge the candidates for best cause/explanation. This, I would argue, is what happened to Dr. Godfrey Louis, and Dr. A. Santhosh Kumar. Dr Louis and Dr Kumar work at the School of Pure and Applied Physics, at Mahatma Gandhi University, in Kerala. Being well versed in their disciplines, and being local to the area, Dr's Louis & Kumar immediately picked on some key elements in the phenomena-premises set and readily applied such background knowledge sets as those of physics; commentary science; and (albeit partially) biochemistry. For instance, picking on the unusual (for that part of the season) thunder and lightning that preceded the red rain, they argued that the sonic boom and fiery atmospheric entry of a meteor would sound and look just like a lonely but very loud thunder and lightning episode. Along the same lines, a low entry angle, in a North-South direction could partly explain the geographic distribution of the red rain. And the red rain itself, you may ask, well -- they continued, those are extraterrestrial, microscopic, life forms.[14] This is not as far-fetched as it sounds. The idea that life on planets can be seeded by life on other planets is a rather well-respected piece of scientific speculation (and thus belongs to any decent background knowledge package in the commentary sciences). Following the nudging of the background packages, Louis and Kumar studied the red rain to find out that, in fact, the red color was not produced by red-colored dust particles. Dust in rain is not an unusual phenomenon, but in this case, the red color of Kerala's rain resulted from the presence of red-brown colored spore like particles. After running some standard tests (indicated by the biochemistry package), Louis and Kumar detected carbon and oxygen, but no bio-molecules (such as DNA). They also claimed that these spore-like cells appeared to be capable of replicating (i.e. make more spore-like cells) at extremely high temperatures. Thus, Louis & Kumar's ITBE could be reconstructed (and summarized) as follows.

P1: Sonic boom-like sound
P2: Entry-like flash lightning
P3: Unknown, red, spore-like shower over Kerala
P4: Commentary Sciences Background Knowledge Package
P5: (*Partial)* Biochemestry Background Knowledge Package
C: A comet bearing alien life broke upon entry showering its contents over Kerala, India.

Please note that I mentioned *two* background knowledge packages, one for commentary sciences, and another one for biochemestry, and that, furthermore, I have noted that the latter package is only been used partially. The reason for this qualification will be obvious in a second, when we go over the other major study on the red rain phenomenon.

As soon as researchers realized that the red rain was caused by a spore-like suspension on common rain water, India's government commissioned two research centers: the Centre for Earth Science Studies; and the Tropical Botanic Garden and Research Institute Tropical, both in Kerala, to produce

[14] Louis, G.; Kumar A.S. (2006). "The red rain phenomenon of Kerala and its possible extraterrestrial origin". Astrophysics and Space Science 302: 175. doi:10.1007/s10509-005-9025-4.

a report on this phenomenon (from now on: the CESS&TBGRIT report).[15] In their report, Sampath, Abraham, Sasi Kumar, Mohanan come up with what they argue to be a more plausible explanation. Unlike Louis & Kumar, these authors brought the background knowledge packages to bear more comprehensively on the phenomena premises. And, in addition, they brought a third background knowledge package: that of the biological sciences. First, they argued that meteoric origins leave a chemical signature on their debris. They reported that this signature was absent from the red rain samples. They also noticed that 90% of the spore-looking elements not only looked like spores but were actual spores, belonging to the genus *Trentepohlia* (the remaining 10% were other, terrestrial microorganisms known to be capable of living in rain droplets). *Trentepohlia* is a lichen-forming alga (lichens are life-forms made up of an alga and a fungus living together). The pigment that this alga uses to store food is red in color, giving its spores a reddish tint that varies in intensity with the amount of pigment stored. When the red rain spore-looking particles were placed in alga growing medium, *Trentepohlia* alga grew diligently out of them. Also, when, a month later, the investigators returned to the area showered by the red rain, *Trentepohlia* lichen were noticed growing everywhere. Even though the authors acknowledge that their findings did not help us understand how the spore bloom got to be siphoned into the rain clouds, they concluded that the most plausible explanation for Kerala's red rain was not a comet delivering extraterrestrial spores but a local weather phenomenon delivering a terrestrial spore bloom. Their argument could be (summarized) and reconstructed as follows.

---

[15] Sampath, S.; Abraham, T. K., Sasi Kumar, V., & Mohanan, C.N. (2001). "Colored Rain: A Report on the Phenomenon." (PDF). *Cess-Pr-114-2001* (Center for Earth Science Studies and Tropical Botanic Garden and Research Institute).

P1: Sonic boom-like sound
P2: Entry-like flash lightning
P3: Red, spore-like shower over Kerala
P 4: *Trentepohlia* spores look like spore-like particles in red rain
P5: *Trentepohlia* grows when red-rain's pore-like particles are cultured
P6: After the red rain incident, *Trentepohlia* lichen grows profusely in area of red rain coverage
P7: Commentary Sciences Background Knowledge Package
P8: Biosciences Background Knowledge Package
C: A local climatological phenomenon rained local lichen spores over Kerala, India.

I hope this short review of the Louis & Kumar analysis, and the CESS&TBGRIT report suffices to illustrate the difference between the two studies. I hope it is now obvious that the most solid the background is; the more reliable our choice of explanation becomes. But now, a most nagging and crucial question remains, how does one determine what will be allowed to count as background knowledge? Or, to put it slightly differently, how do we know what makes a background knowledge package solid?

4.2.3 What should count as Background Knowledge?

This is not an easy question. As we have seen, different background knowledge packages will result in different "best" explanations being chosen. If we do not fix this problem, then choosing your background knowledge (and thus your "best" explanations) boils down to a matter of arbitrary, personal preference; a mere subjective choice. But we embarked on this journey seeking objective ways to ground for facts, so this problem must be resolved: bad explanations can, rather literally, *burn* you.

4.2.3.1 Explaining Witches

Imagine you were unlucky enough to be born an ambitious, independent 1500's Eurpean woman. It's quite likely that you would have made quite a few enemies pretty soon and that, also sooner than later, one of them would have conveniently decided that you were a witch and reported you to the Inquisition authorities. The Inquisition authorities would have, also very likely, read a very popular handbook in their trade: *The Malleus*

Maleficarum (The Malleus, form now on). Two inquisitors for the Catholic Church wrote this book in 1486: Heinrich Kramer and Jacob Sprenger.[16] Kramer and Sprenger piece gives us an eerie window into how inferring to the best explanation can sometimes go very, very awry.

The Malleus was choking full of instructions on how to spot and deal with witches and other Satanic phenomena. As mentioned in passing above, the book also included several explanations of these phenomena. It explained, for instance, why women (and not men) are likely to become lured by Satan into witchcraft. Talking the biblical story literally, *The Malleus* points out that God created Eve, the first woman and mother to all the rest, from Adam's rib. These authors cunningly point out that, quite obviously, the rib is not a *straight* bone; it is *bent.* Do you get it already? Women, according to the *The Malleus*, were *bent*; a *crooked* creature from the beginning!

I hope you realize that this is a ridiculous explanation. The troublemaker here, quite obviously, is the background knowledge package. There are way too many problems with it for us to deal with all of them, but let us outline two of the silliest and most crucial ones. First, these authors are requiring a literal interpretation of a text (*The Old Testament*) that should be subjected, instead, to a religious one.[17] Second, they require the idiosyncrasies of one particular branch of languages to have had a determinant effect on the biology and psychology of all of humans (and this, furthermore, needed to have happened before this particular language set had ever arisen!). In other words, Kramer & Sprenger argue that the double meaning of the word bent and crooked, in the Germanic languages, must be mirrored by the stuff to which they apply. So that if something happens to be amenable to be described by a Germanic language *speaker* as *bent* or *crooked*, then the nature of this something will be also bent or

---

[16] You can read the entire (albeit in Latin) text online at: http://digital.library.cornell.edu/cgi/t/text/text-idx?c=witch;idno=wit060. Otherwise, there is a good English translation: Montague Summers. *The Malleus Maleficarum of Heinrich Kramer and James Sprenger.* Summer Dover Publications. 1971.

[17] This is, of course, not my personal opinion: it has been the Vatican's official position, for at least the last 70 years or so (I'm referring here to the re-foundation of the Pontifical Academy of Sciences, by Pius XI, 1936. For more on this see, for instance, Pope John Paul II address to the Pontifical Academy of Sciences (October 22, 1996), at http://www.newadvent.org/library/docs_jp02tc.htm.

crooked. Geometrical arcs, for Germanic language speakers, are devious, crooked figures. (Silly.)

But, you might say, this is silly *for us*, but why should it have also been silly *for them*? This is a very good question. The problem of subjectivity creeps up yet again. If ITBE's are to be an objective piece of thinking, then this type of thinking should work regardless of where in history it is placed. To put it slightly differently, we should be able to argue that the conclusion above is not a good explanation, even if we were immersed in the same time and place as Kramer & Sprenger.

4.2.3.2 May the Force (of Reason) Be With You

4.2.3.2.1 Have we gotten any certainty yet?

No. We are almost done with our assessment of how ITBE's work and, just like in the case of inductive moves; we should already be able to conclude that there is no certainty here either. Again, we are dealing with a conclusion that lies *outside* of the set of its premises. In this case, unlike the case of inductive moves, the conclusion is not the premises' generalization, but a *cause* that, when proposed in the conclusion, turns the phenomena premises (via permission of the background knowledge packages that apply) into its *effects*. So, phenomena premises and background packages can all be true and still (even when dealing with the most excellent abductive moves) ITBE's cannot guaranty the truth-value of their conclusions. Deductive tests applied to the best explanations produced by ITBE moves, as we have also seen with the case of inductive moves, can support and give us justified hope, can refute the bad explanations with certainty, but never prove them to be true.

4.2.3.2.2 Plausibility: in depth

The only way to escape subjectivity is by letting the facts, and not our personal preferences decide which is the best explanation. And by this I do not mean something like "time will tell." We saw that facts are always the result of a piece of (excellent) reasoning. Since we now understand that excellent (i.e. most objective) reasoning only comes in three modes: induction, deduction; and abduction (*IDA*, from now on).[18] Consequently,

[18] From now onwards, and for the sake of brevity, we will always refer to the compilation and combination of all of our best inductions, deductions, and

requiring background knowledge packages to be made up *exclusively of what IDA says*, should be one and the same as ensuring that they are made up *exclusively of most solid facts*.

We saw above, when we were studying how we tackled the red rain event, that the background knowledge package was *allowing* some explanations a possible candidates while denying that status to others. *The explanation* thunderstorm was denied the status of best explanation as soon as the red (not just normal) rain premise was brought into our thinking. Similarly, as soon as the authors of the CESS&TBGRIT report properly categorized the red rain event as a local, climate-biological phenomenon, and then brought to bear all appropriate (that is, according to IDA) background knowledge packages on it, unfitting explanations were disclosed and discarded (e.g. the one about a comet bearing alien life forms), and best explanations were provided -- when available (a local lichen bloom spread by a still poorly understood climatological event – no good explanation for the latter).[19] Thus, deciding whether or not an explanation is the best one boils down to assessing its fit with the established relevant facts that make up the background knowledge package. To be a bit more precise and to express this idea succinctly, from now on, the presence of a fit between a would-be-best explanation and its relevant background knowledge packages will be addressed by the word *plausibility*. And we now understand that plausibility is one and the same with what IDA says. In other words, we will say that an explanation is the most *plausible* when it best fits all relevant and most excellent background knowledge packages.

Can we now discard *The Malleus'* explanation? Can we refute the claim (as we should) that your mom, sister, daughter, and or girlfriend are all *bent* and *crooked* by origin and nature? As we concluded above, background knowledge packages must be made up of only excellent pieces of reasoning. We identified two crucial pieces of thinking in Kramer & Sprenger argument that were sort to speak rotten. The first problem was their allowance of a

---

abductions as *IDA* (an acronym made of the first letter of each kind of objective reasoning and pronounced as the woman's name). So, what our best reasoning says is going to be talked about as *what IDA says*.

[19] There is another important lesson to learn here. The background knowledge package, sometimes, deems that there are no good explanations out there. In these cases, satisfying our need to explain by patching in some implausible explanation is, of course, a very bad idea. Explaining away *is not* the same as explaining period.

religious text (which, by definition, is *not* the result of IDA but of revelation) to be part of the background knowledge package (which, as we have argued, should be exclusively an IDA set). The second problem was the claim that the nature of language determines the nature of the physical world (or at least part of it: human biology). Consequently, and applying the correct background knowledge packages, we can see that there are no IDA reasons to believe that the human female was literally created from a male bent bone; and that there are no IDA reasons to believe that Germanic languages determine the nature of physical properties.

However, some might still complain, is unfair to require Kramer and Sprenger to have had access to modern background knowledge packages. And that would be a very sensible point to make. Nonetheless, as we briefly discussed in our previous footnote, sometimes IDA has nothing to say. SO, even though these authors did not have access to modern IDA conclusions, they should have listen to their own time's IDA recommendation. Even in the 1500's there were enough IDA support to convince anybody who would care to be convinced, that there were no objective reasons that could fit the witches explanations. Instead, what IDA would have said then should have been something along the lines of "restraint your subjective urges and keep researching the subject objectively."

Alas, I know that, most likely, Kramer & Spranger (and everybody else that enjoyed getting their way at whatever cost possible) would have not been moved by these reasons. Not everybody cares about what IDA says. But noticing the force of reason and caring about it are two very different things. And promoting the second one is well beyond the scope of this treatment. What we are trying to do here, quite humbly, is simply to help you notice. Surely, one can always hope that once you do, then the caring would likely follow. If you think about it, there are actually good reasons for hope.

Reason has its own force, regardless of the subjects' idiosyncrasies. This is sometimes easy to overlook, as an old, one peso coin, from Chile Illustrates (see Fig. 4-1).

Figure 4-1. An old Chilean, one peso coin that reads, on one of its sides "By Reason Or By Force."

Except for very few circumstances, we have as little control over what is reasonable (and what is not), as we do over what we can perceive or not. Let me demonstrate this with the help of a little thought experiment. Imagine I can install a little window in your head that shows me exactly what you are thinking (promise it wont hurt!). Now, I put my hand up and ask you to raise your eyes to look at it. If you look, can you help but see that my hand is up? Of course: not. Imagine now that, apparently not knowing much about human nature I proceed to tell you that I'll pay you ten dollars if you do not see it. To my (ignorant) surprise, even though you agree to do your best (let us imagine that you are very greedy), I see through the little window in your head, that you are still seeing my hand up. Frustrated, I offer more money: a million! I scream. When I realize that you are still seeing my head I pull a gun and threat you not to do so (sorry). Of course, I does not matter what I do; if your eyes are open and you are looking in the right direction, you *cannot help but* to see the hand. The force of the world and our perception easily overwhelms our will. You should not be surprised to realize that the same would be the case if I were to command you believe that two plus two equals six; or to conclude that if all N's have property Φ, and n is an N, then n *does not have* property Φ; or that if A is greater than B, and B is greater than C, then A is greater than C. The point being that you have as little control over your perception as you do over what makes sense, or not.

So the Chileans who created this coin, I claim, were wrong. Their message is redundant: once you point someone's mind towards reason, they have no choice but to see it, and physical force should (hopefully be) redundant.

Now, as we have already established above, whether one cares or not about *following* what reason dictates is another matter completely.

4.3 Semi-Final Wrap-Up

Let us go back to Gliese 518g but jumping ahead, about 5 years after the faithful creation of the committee on fact-finding recipes. Our decedents have purified each of their four methods and realized that they were dealing with just three different kinds of thinking: inductions to discover patterns or correlations; deductions to pump fact from other facts and to perform tests; and abductions to discover causes. They also understand that the word *fact* was just too presumptuous. Their correlations can aspire to be super-strong, their causes superbly plausible, their tests incredibly well confirmed, and their fact pumping (albeit form incredibly well supported -- but never certain -- other facts) indubitable. But the idea of factual certainty flew out the window as soon as understanding came in and settled.

A major turning point in Gliese's fact-finding enterprise happened when a second, subsidiary committee was formed once they hit upon the idea of *plausibility*. As we ourselves discovered, abductions are all about finding the most *plausible* explanations, and plausibility is all about what IDA says. But, both Glieseans and us know that IDA cannot really *say* anything because IDA is the objective output of our best inductions, deductions, and abductions -- and precisely *not* a (subjective) person. So their IDA Committee (that's how they called their subsidiary to their Fact-Finding Committee) set out to devise a way of having IDA -- not another committee member -- talk.

The IDA Committee soon found out that individual Glieseans could never speak for IDA individually. First, individuals have personal biases that are not easy to keep at bay. So it was not always easy to make them let go of their favorite causes and correlations, or to accept their favorite would-be fact refutation. Second, and of equal importance, no single Gliesean could store and apply all the IDA results, even when they were talking about (what had originally started as) a small specialized field.

So, the IDA Committee invented what here on Earth we call expertise and peer review. The process of peer review was effected by a society of experts in every active field of fact-finding. To be an expert in any given field, all any a Gliesean native had to do was to stick to concluding stuff only via induction, deduction, and abduction, and writing up a report of their findings that clearly specified how any other expert (or for that matter, anybody else who cared to do so) could, still only by endorsing IDA, reach exactly the same conclusions. Thus, what IDA said emerged, not in a mysterious way, but as a direct result of the accumulation of these reports (or, more precisely, of those that survived the scrutiny of other experts).

I hope that the Gliesean story is ringing all sorts of bells inside your head. It should sound very, very familiar because, here on Earth, we converged to an identical fact-finding institution: we call it *science*, and to it, we will move next.

5 CHAPTER 5 -- SCIENCE: FACTS ON STEROIDS

5.1 A Few Words on "Understanding"

When one thinks about understanding, it seems reasonable to think that there must be many different ways to understand the world. But here, I will propose a more precise meaning of the word *understand* that will bar this ambiguity. Typically, *understanding* is what follows when we get an *explanation* and, as we have already seen, explanations are what *causes* provide. Thus, one could conclude (probably a bit hastily) that "understanding" is the word we use to describe the psychological state that follows receiving an explanation. But this is a *subjective state of mind*, and as you already know, not necessarily very informative about our concern: states of affairs, or of the world – instead of the mind. So, in this essay, when we say "understanding," we will be referring to the grasping of true causes: the understanding of the *causal network that lies behind the phenomena* under consideration.

"Grasping" might sound a bit vague, but what we mean here by it is the process of getting something like a map that will take us from the responsible causes, to the phenomena that they produce as their *effects*. So, for example, we will say that we *understand* how cigarette smoking causes cancer when we grasp the detailed maps of the causal networks that take a person from puffing to developing cancer.[20]

Scientists are after such understanding, and to find it, they deploy precisely the same every-day fact-finding maneuvers that we discovered (and recognized as our own) with the Glieseans. But they apply these strategies to their purest distillate form and maximum degree of effectiveness possible. To the understanding of this purification and maximization of our every-day, fact-finding methods we will move next.

5.2 A Patterned World

Science usually starts and ends with patterns (usually called phenomena or regularities of nature). Phenomena are grasped, mainly, as a result of

[20] A caveat: we are talking exclusively about *factual* understanding. One might wish to argue that there might be kinds of non-factual understanding. Regardless of what non-factual understanding might entail (if anything at all), here, we are only concerned with facts. And since we now know that facts are what IDA says, then factual understanding must also be what IDA says.

inductive generalizations. Some generalizations seem easy to get. Some phenomena seem to stare at you right in the face. Or, to put it slightly differently, we do not seem to need sophisticated inductive maneuvers to spot some regularities. The sun coming up each morning (albeit sometimes behind a cover of clouds); the tides; or the seasons are all patterns in nature that are hard to miss. But, of course, this is not the case for every generalization of pattern out there. Furthermore, there is always the possibility that we are jumping to a generalization too quickly and rather naturally.

Many times, to avoid such pitfalls, we enlist the help of powerful inductive/statistical search engines. Scientists collect large amounts of possibly related variables and unleash such statistical hounds onto them. This sophisticated inductive machines comb the collected data, sniffing out correlations that might otherwise go unnoticed. We already visited one such a case when we discussed the correlation between shaving and coronary disease.

But, you may ask, *where* do scientists unleash these powerful inductive programs? Or, to put it slightly differently, how do scientists know where to look for possibly correlated variables? Would it be possible to just collect data blindly and find them? Not really. Even though we have separated induction, deduction, and abduction to understand each one of these modes of reaching conclusions on their own grounds, these three kinds of reasoning usually work as members of the same pack, complementing each other, like The Three Musketeers. Consequently, hunting for inductive correlations is always far from an entirely blind search. Accumulated scientific facts (of course via IDA) nudge scientists to look this or that way, to suspect that a correlation might lay here or there. For instance, let us look at how our understanding of the mechanics of evolutionary change nudges us to look for very specific patterns and correlations in nature.

Once we know that life in this planet is the result of an evolutionary process, we can deductively conclude that there must be very precise patterns out there that bear the stamp of to the forces behind it. As we saw when we used deduction as a fact-pump, to prove that natural selection (the evolutionary force that builds adaptive change) was an inescapable outcome of the way the world is, we learnt that species diverge by accumulating tiny heritable changes. We understood that when these changes gave creatures an edge in their local struggle for existence, natural selection would end up accumulating what we can call *adaptive* or functional design – like eyes, whose adaptive function is to send light reflections to a central processor (the brain).

Now, once we understand how the causes of evolution work, one must also deduce that natural selection is a very peculiar kind of designer. It can only design by tinkering; never from scratch. All functional structures designed by natural selection (such as eyes; stomachs; or wings) cannot be designed in the same way an intelligent engineer would. Naturally, the design has to be "intelligent" in the sense that it must work: serve the bearer in the local struggle to produce viable offspring, *but nothing more.*

So, more precisely now, natural selection designs with at least two severe constraints in "mind."[21] Number one: adaptations are not designed from scratch: they are always tinkered form some previous structure (that could be an adaptation for something else) that is gradually co-opted to be put to a novel use. Number two: natural selection's adaptive design will be only as good as the local struggle requires, nothing more.

So, armed with these IDA facts, scientists are directed to go out and check for patterns in promising places. Thus a sometimes positive and sometimes negative feedback loop is set in motion. Established IDA facts make scientist look for patterns in promising areas, and what they find out, then feeds further IDA thinking that will support, refine or discard (whichever IDA so demands) the causal landscape they are investigating. Let us see three such guided pattern searches to illustrate this point.

5.2.1 Just One Way to Skin a Banana

Since natural selection proceeds by the accumulation of tiny changes, we should, for instance, be able to trace all life to just one or a few original ancestors. Along the same line of reasoning, the necessary processes of life should be highly conservative. In other words, we shouldn't find too many reinventions of the wheel when it comes to fundamental things such as breathing, breaking down food, or growing. Putting it simply, the fundamental recipe for making, let us say, a banana, should not be that different from that for making a human.

[21] So to speak, of course, because natural selection is just a mechanism that follows from the way things are out there, it has no purpose, or mind whatsoever.

All forms of life in this planet reflect this pattern. As expected, we share about 50% of our genome (our species typical full dose of DNA[22]) with that of a banana; and over 98% with that of our most recent close relatives: the chimpanzees.[23]

5.2.2 Tinkering Only

Since natural selection is a tinkerer, there must be tons of structures that have been co-opted into new uses (remember, no building from scratch for natural selection). Birds offer a feathery example. Dinosaurs did not go extinct, that is, at least not entirely. One of their branches, present birds, became very widespread and diversified. There are very good (IDA) reasons to believe that dinosaurs, of the pre-bird kind, were already covered in feathers.[24] But this (pre-bird-stage) dinosaur feathers' main purpose was not that of flying: the main function of dinosaurian fine feathery coverage was thermoregulation (think about the goose-down that fills our excellently insulating down jackets) and sexual advertising (like in present day – and most generally -- male birds do, while displaying their colorful plumage to attract the females). (The next Jurassic-Park kind of movie will need a bigger budget to accommodate the extra CG that this new finding will require.)

---

[22] DNA is the molecule responsible for passing on hereditary information. In the sequence of its building blocks, the recipe of how to make a fully functional adult is (so to speak) *written*. Also check the following footnote for more resources on DNA's structure and function.

[23] See, for instance, PBS' NOVA "Cracking the code of Life," at http://www.pbs.org/wgbh/nova/genome/program.html.

[24] See, for instance, the New Scientist's article: "Were all dinosaurs beasts of a feather?" 19 March 2009, available at http://www.newscientist.com/article/mg20127005.100-were-all-dinosaurs-beasts-of-a-feather.html.

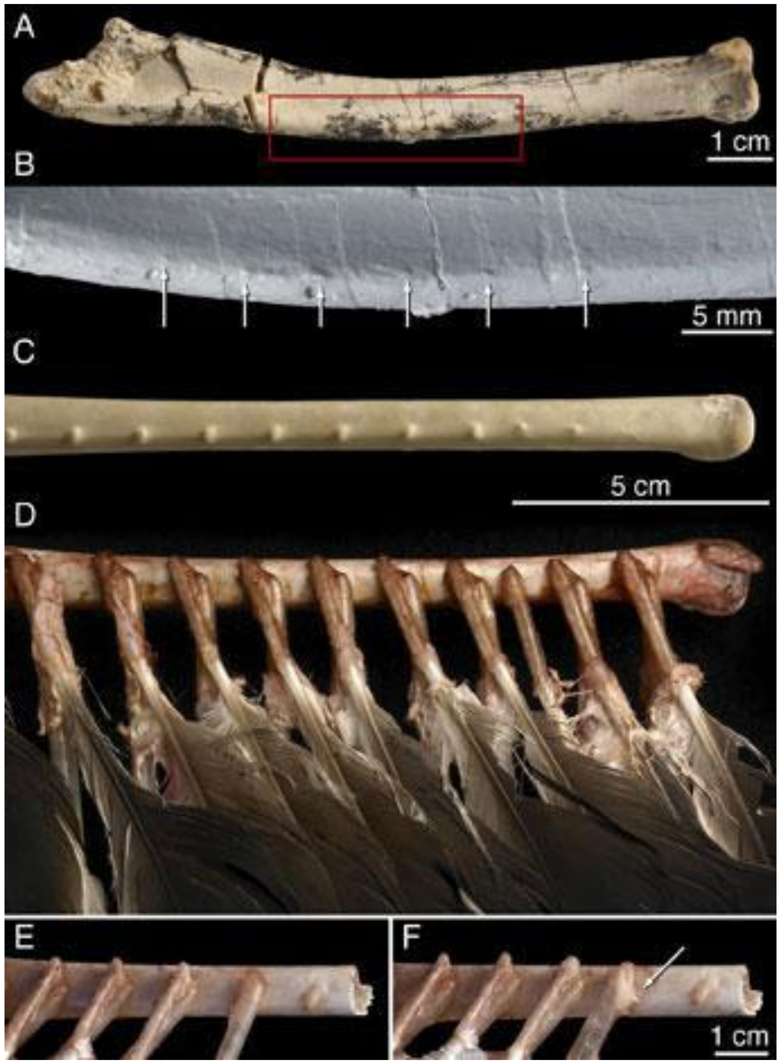


Figure 5-1. The little knobs on the *Velociraptor* fossilized bone (top two pictures in the figure) look much like those in the modern turkey vulture bone (four bottom pictures in the figure) (Image: *Science* -- (DOI: 10.1126/science.1145076).

5.2.3 Unintelligent and sometimes quite nasty design

As we have seen, natural selection can do nothing but tinker (no redesign from scratch available), and its criterion in designing excellence is, to say the least, surprisingly narrow and mundane. More precisely: a given design is selected if it allows the struggler to out-reproduce the other strugglers in their local province. This is bad news if one is at all concerned with the quality of adaptive design that this process produces. It should now be obvious that natural selection's design does not have to be intelligently engineered; or moral; or displaying any sort of foresight (to mention just a few of this process' lacking's). In short, natural selection cannot design like an intelligent, kind, or concerned-with-the-future kind of designer. As long as the bearer of the design in question out-reproduces its neighboring strugglers, natural selection will have no qualms perfecting adaptations that might, in the long run, and for instance, destroy the species in question. To think of an example of this, all you have to do is to look at the countless species of parasites that drove their hosts, and themselves (ups!) to extinction. Or to realize that only an infinitesimal portion of all the species that ever existed are alive, with us, today.

Not only does natural selection lack foresight, it has not even a myopic concern with present quality and utility either. Natural selection's artifacts are of rather suboptimal quality (that is: in smart engineering terms). One does not have to go far to see this (no pun intended!). Vertebrate eyes (and that includes human eyes, of course) bear the signs of their lowly origin imprinted in their design specifications. As you might know, the retina in our eyes works as a kind of screen receptor, capturing the light that bounces off the objects we would like to *see*. The retina's previous usage was not very friendly to the new requirements for higher processing and its associated rewiring. Incapable of redesigning form scratch, but always mindlessly undaunted, natural selection diffused the wiring *over* the retina, punched a hole in the middle of it, and fed all necessary wiring to the brain. Of course, the hole in the middle of the receptor screen is not a very bright idea -- in engineering terms. No engineer (not even a dumb one) would have created such a blind spot, obviously compromising the functioning of the light receptor. Even a mediocre designer would have re-wired the retina so that the wiring was *behind* the receptor and not *in front* of it – no need to punch a hole through it now: duh! Other organisms, like octopi, or squids, lucked out in the design lottery and got, by shear evolutionary luck of the draw, a better design than us vertebrates.[25]

Similarly, the human body suffers of tons of highly predictable malfunctions that could have been easily avoidable by a dimly intelligent designer. Again, due to a variety of ancestral-baggage reasons, we have things such as easily detachable retinas; blocked male urethras; and a tendency to choke to death (to mention just a few).[26] More precisely, the same bad wiring that

---

[25] For more of this, you can go, for instance, to PBS' NOVA "Evolution," at http://www.pbs.org/wgbh/evolution/change/grand/page05.html.

[26] For a good, accessible read on this one you can go to Scientific American's

we mentioned above (form the front of the retina instead of the back), is largely responsible for retinas that detach way too easily. A prostate gland that *wraps around* the urethra (instead of simply laying beside it) tends to constrict it as soon as the prostate gland becomes inflamed (which happens to one out of every two men – a solid inductive move, by the way). Finally, making a talking human out of a non-talking human-chimp common ancestor involved several unfortunate compromises. It required, for instance, breathing and eating sharply bent pipes that are poorly separated -- and that any good engineer would have simply redesigned. According to the Home Safety Council, death by choking and suffocation is the fourth home-injury cause of death in the US.[27]

I also called natural selection's designs *nasty* (in the last main heading) and *amoral* (in the text). The point I am trying to make here is that natural selection is not only unintelligent, but that it also designs with an utter disregard for desirability or any kind of etiquette. To establish this, two examples should suffice. Parasitism is the most common way of life in this planet.[28] And some of these parasites are really nasty. Think, for instance, about the Amazonian *candirú*.[29] The candirú is a little parasitic fish that *usually* lives stuck to the gills of other, bigger fish. The candirú has piercing barbs in its head that allow it to hook itself to the gill of its host and to feed of the bleeding that this crude lodging mechanism induces. This parasitic fish finds its hosts by the trail of urea that the gills produce. Unfortunately (for those mammals that enjoy peeing in the Amazon basin), urine (as its name aptly indicates) also has urea. The candirú's urea detectors make it very capable of following the scent of urine, of swimming up the urethra of its relaxing host, anchoring itself to its walls, and living off the blood that slowly oozes out of the injuries that its barbs cause. Reality, once again, outperforms the goriest of horror films!

"If humans were built to last Built to Last;" by S. Jay Olshansky, Bruce A. Carnes and Robert N. Butle March 2001

[27] http://www.homesafetycouncil.org/AboutUs/Research/re_sohs_w007.asp

[28] Windsor DA (1998) Most of the species on Earth are parasites. Int J Parasitol 28:1939 – 1941.

[29] If you don't believe me, see, for instance, "candiru." Encyclopædia Britannica. 2009. Encyclopædia Britannica Online. 03 Oct. 2009 <**http://www.britannica.com/EBchecked/topic/92428/candiru**>.

Sorry for the unpleasantness that you must be feeling now, but I hope this suffices to illustrate the search of IDA informed phenomena (like those of unintelligent and nasty design in nature). Let us now move on to less disturbing arenas (I promise to chose lees disturbing examples, form now on). As we discussed above, factual understanding is about getting two things right: the patterned effects, or phenomena, and the patterned causal networks behind them. We should now have grasped how scientists discover the first. Now, we will move to analyzing how do they get the second. To do that we will now bring into our discussion the quintessential set of scientific products: scientific theories; models; hypotheses; and experiments.

5.3 The Works: From Theories to Experiments and Back Again

At the beginning of this essay, when we set out to discover our humble, every-day fact-finding methodologies, we saw (with the all-too willing "help" of Gliesean Tarantulas) that *the Trial & Error Method* figured prominently in our list of most-trusted tools. When we put this method under scrutiny, we realized two things – one of them rather terrible. First, that this method amounts to nothing more, nor nothing less, than a deductive test. Second -- the terrible one, that tests cannot possibly prove anything. A positive test result leaves room for doubt even when it is derived from true premises. Putting it shortly, since the positive outcome of a test amounts to an invalid deduction, it cannot possibly work as a fact pump. When it came to testing, the only fact pump we found was on the negative-results side: we can prove our hypothetical statement (the one being tested) wrong, but not right (we called the latter a "corroboration," or "confirmation," and the first "refutation").

Now, just as we, in our every-day life utilize testing to corroborate, or to prove wrong our ideas about the world, scientists use it to test their suspected patterns to either corroborate, or prove them wrong them. But because of the way this is done, scientific corroboration and refutation is far more carefulness and precise. To understand this, we must talk about supermodels.

5.3.1 Abducing & Testing Supermodels

The word *model* in colloquial English has many meanings. Here, we will restrict this ambiguity a lot, but let us start with one aspect in the colloquial usage that we will keep. The word *model* bears the notions of an ideal version of some real thing. An *ideal* version of something is a version that captures some salient feature of the real thing while leaving others behind. Thus when we talk about a *model airplane*, we mean to say that we have a small construct that captures the overall outside proportions of an aircraft, but that it leaves unrepresented its true size. Similarly, one could argue that the vernacular use of the word *supermodel* also reflects the fashionistas' attempts to capture their ideal proportions of the feminine figure leaving the actual feminine averages aside. Scientific models could be thought of as supermodels, but in a very different sense of the word "super." Here, the "super" does not refer to some exaggerated property, but to the fidelity with which the model captures what is supposed to be modeling.

Scientific models are idealized versions of the world, but what they are trying to capture is not necessarily the world's appearance (like, let us say, in a modeled airplane), but its *causal structure*. To understand the subtleties entailed here, R. G. Giere's idea of scientific models as *maps* is very useful.[30] Giere proposes we understand scientific models as *causal* maps of the world. One of the key advantages of understanding models as maps is that it takes us away from the misleading idea of scientific models as somehow *picturing* the world. As Giere points out, maps, unlike pictures, do not attempt to *represent* whatever it is they are mapping; they only intend to capture some salient feature or other. So, a topographic map, for instance, tries to capture the relief and relative distances between relief points of a particular area, and it does so with a specified scale. Similarly, a street map, tries to capture the relative position and distances amongst the streets of a certain area (and not, for instance, the color of the buildings). A road atlas tries to capture the same, but amongst major roads (without usually bothering with the precise width of the road, for instance).

The *capturing* that scientific models do is exclusively *relational*. But unlike a street map (that captures relational *distances* between streets), a scientific map captures how different *causal networks relate to each other and their effects*.

---

[30] (This book is also an excellent further-reading suggestion to better understand the nature of scientific thinking in general.) Ronald N. Giere. "Understanding Scientific Reasoining" 4th ed. (1997) Thompson-Wadsworth.

5.3.2 Laying down The Laws

When people talk about science, the word *law* figures prominently in their discussion. Now, as we begin to understand what scientific models are, we can get a better grip on the nature of this useful term. Just like legal laws, *laws of nature* are those general, or universal rules that *regulate* the natural world. The idea of regulation is key here.

First, we are not talking about mere generalizations here. Or, putting it a bit more precisely, not every correlation (the kind a good inductive move would produce) should be considered a *law of nature*. Because, secondly, and unlike a mere correlation, laws of nature *make things happen* -- they don't *just happen*. Let me illustrate this with a rather humble example. Probably, there is a correlation out there that truthfully describes that, if you find a fridge, anywhere in this planet, that fridge will most likely be white. So, *it just so happens* that Terran fridges and white fridge paint are correlated. But, obviously, this *must* not be the case; there is no *law of nature* that *determines or regulates that this correlation must happen.* A key question that should be probably popping in your mind right about now is: what gives laws of nature their regulatory powers?

The regulatory aspect of laws of nature simply refers to their causal powers. When we talk about laws of nature we are really talking about the particular rules that determine the particular behavior of a certain set of causes and their certain effects. So, since as we have seen, scientific models map scientific causes and their effects, laws of nature are the rules and stipulations that map out the relationship amongst causes, relevant conditions, and effects. For the sake of precision (more on this below), these relationships are usually expressed in a mathematical form.

Think, for instance, about Einstein's famous equation $e=m*c^2$ (or: the total energy of a system equals its mass, times the velocity of light squared). Einstein produced this law of nature, after modeling, with deductive precision (e.g. mathematically) the rules governing the causal network that involved things such as space-time, mass, and energy. Once this network was modeled, Einstein was able to deduce, for instance, that tiny, tiny bits of mass could translate into tons, and tons of energy (a realization that underlies such terrible things as atomic bombs; but also nuclear plants). Or, that you should not even try to move at the speed of light because, unavoidably, your mass would approach infinity (and it is impossible to move if you are that big).

You probably heard some of this before. And I bet you that the word *theory* must have been floating around quite prominently in those conversations. Now that we have an idea of what a model is and what laws of nature are, we are ready to tackle the idea of a *scientific theory.*

5.3.3 Scientific Theories

As it should now be becoming obvious, scientific models do not exist or work in isolation. As explained above, scientists devise models to map what is going on (at a causal level, that is) behind the curtains of the phenomena we can observe. And everybody knows that, in this world, many causes work in concert to produce these regularities of nature, or phenomena. So, since scientific models map nature's causes, and since nature's causes work in complicated networks, scientific models must work in packs, just like the causes they are mapping. Scientific theories are packs, or families of interacting models that tackle, in a concerted effort, a particular chunk of the natural world.

Let us put this fresh understanding to work to dispel a common misunderstanding out there. Imagine that someone claims that this or that observation refutes, let us say, the theory of evolution (I bet that you have heard this one before). If you understand what scientific theories are (complexly integrated packs of models), you know already that this claim makes no sense at all. You should now understand two things. First, as we learnt in the chapters before, refutation is a logical maneuver of surgical precision. Second, as we learnt just now, scientific theories are complex families of interconnected models. Refuting such a beast with this or that observation, should sound sillier than the prospects of killing an elephant with a single mosquito bite. Actually, the refutation claim is even sillier, because in the case of a scientific theory, unlike in the case of an elephant, you would be at a loss trying to find a heart, or a brain, or any vital-organ-like structure! Like with any other decentralized, complex set of integrated elements, an instance of pinpoint damage is certainly cause of adaptive modification, but it cannot in itself, singlehandedly justify the abandonment of the whole.

Now, we are ready to tackle what I loosely just called "adaptive modification;" referring to the aftermath and repercussions of a refutation instance. Let us bring the analysis on with the following question. If refutations cannot kill scientific theories, what is it that scientists refute, or confirm, and how can scientific theories get better via such processes?

5.3.4 Testing – the surround version

Let us think, for instance, about *genes* – the scientific model that maps one of the central causal agents behind heredity.[31] Let us picture a sexually reproducing creature. A *gene* can be modeled as any piece of DNA that triggers some crucial difference in the way this creature is, and is usually not broken up when the creature produces its sexual cells (sperm cells, for males, and ova, for females), or during the reshuffling that takes place when these cells join to make an egg (that will generate, in time, and all other things being equal, an embryo).

Now, as we have seen, scientific models never work in isolation. Modeling genes requires the integration and articulation of a set of models within the theory of genetics. But the articulation does not stop there, as we will now see, before and after deriving hypotheses to test a given model, contextual (to call them somehow) models become integrated from neighboring, and not that neighboring disciplines.

To map the causal network of what a gene is, and what a gene does, one needs causal maps coming from disciplines that lay below, above, and sideways (so to speak) to those that pertain to the immediate causal neighborhood of the gene and gene-effects in question. This contextual disciplines range from physics and chemistry, to evolutionary and population biology (to mention just a few of the neighbors). To put it in slightly different terms, and a bit more concretely, a full blown model of how genes work would require pieces of understanding coming from all sorts of different directions. Obviously, what genes are made of should mater greatly. So one would need to understand the biophysics and chemistry of DNA and its entourage of biochemicals (the molecules and processes that, as we mentioned above, codes and bears the information that is passed from one generation to the next). To understand all aspects of a gene's functionality and nature, it would be necessary to understand the original how and why of genes; so one would have to bring to bear neighboring models from evolutionary biology, and developmental biology. And this application of progressively more and more distant models is only getting started. As it should be easy to see, to figure out, let us say, how genes move through space and time, one would also need to bring to bear, for instance, the laws of their bearers' mating strategies; and their population

[31] For a good, online learning resource on genetics in general you can go to: http://learn.genetics.utah.edu/.

dynamics. In addition, when, for instance, the gene model is applied to human populations, further models must be brought into play -- such as those belonging to the social sciences, medicine, and economics (to mention just a few). This ever-expanding network of interacting models only gets exacerbated when one performs laboratory experiments.

When one tests a model in the laboratory, one needs to bring to bear an array of even more foreign sets of models that are necessary to design, and perform the laboratory protocols. For example, think about the set of contextual, or auxiliary models (again, to call them in yet another way) that will be required to design methodological and statistical strategies of data collection and analysis. Or (bringing to bear yet more vastly different disciplines), consider the models that are necessary to design and utilize whatever laboratory apparatuses the experimenter might require.

The refutation-that-kills-an-entire-theory claim must be sounding sillier and sillier by the moment. It should now be easy to see that, whenever a refutation happens, a great deal of time is spent by the researchers trying to figure out which part, or parts, of which model, or models, in the expanded, articulated set, was or were responsible for the negative results. But there is an even more substantial message to take home from this exploration. *The wide context of models that are brought to bear any time a given model is put to work, or to the test, amounts, in itself, to an additional IDA check on the model in question.* Consequently, and even before we start thinking about the nature of experimental tests (immediately below), we can see that models are, sort to speak, under constant fire, and from all corners. More precisely, this contextual "testing" comes from the model's success (or failure) to articulate with all well established models that surround it. Consequently, *before* a given model is put to the test via actual experiments, it success, or failure at integration with the network of relevant, contextual, well established models amounts, in itself, to a test for the newcomer.

To get a better grip on this additional form of testing, it might help to think of an analogy. Imagine that somebody suddenly claims that there has been an extra person living in your house. Well, she better be able to provide some evidence. If you think about it for a second, what this "evidence" amounts to would be nothing more, but nothing less, than IDA demands for articulation with everything else that has happened in your house this far. Naturally, and also, if this mysterious person can help explain some previous happenings in the house that remained unexplained by what was known about the house's inhabitants (let us say, an extra bed that appears to be in use, missing food, noises in the night, etc.), even better! Similarly, when a new scientific model is proposed, the first hurdle it encounters is the requirement of an exquisite fit with its context of well-established models, and the novel explanation of otherwise problematic or disconnected facts. If the new model does not fit the established context, then most scientists naturally think that there must a problem with the newcomer. Of course, in some extraordinary cases, this thinking is abandoned when the newcomer does an even better job than the older, established locals. But in any case, this contextual articulation is not the be all and end all of it, but only an invitation for actual, experimental testing. So, whenever a scientist discovers a model that fits everything else we know, and then explains some, she/he has the right to get very excited, and to start thinking about performing some experimental tests.

We can call this type of model support "support from above." And now, we are ready to consider "support from below:" the kind that is provided by a deductive test (of the Gliesean Trail-and-Error type). Now, we are ready to tackle the question that we posed at the end of the previous section: what is it exactly that scientists refute or confirm? Understanding what happens when scientist draw hypotheses, predictions, and experiments will finally allow us to fully understand one of the crucial roles that deductive moves play in science.

5.3.5 Testing by Precisely Deriving Predictions

The terms *hypothesis* and *prediction* have a history of abuse. Just like with the case of "theory," they have become mainstream and as such have accumulated enough meanings to make them rather ambiguous and thus unhelpful. Here, we will define them in ways that highlight and explain the crucial (deductive) role that these key elements play in science.

A hypothesis is a *claim*. It is the proposition that a certain scientific model fits, or maps a certain part of the causal world. Remember that this, in broad strokes, was the very reason why models are built. As you can imagine, claiming that the model fits the world is a big claim and it better brings some substance behind it. The weight, or justification behind this ambitious claim is brought by a deductive maneuver.

The logical, or deductive relation that hypotheses exploit is that of the *if-X-then-Y* nature. Let us call any statement of this nature a *conditional statement*. For instance: "If you are in Alaska then you are in North America" is a conditional statement. In a hypothesis, the *if* part refers to a proposed model, and the *then part deduces* from it a very *precise state of affairs.* In other words, a hypothesis claims that*: if the model in question is true, then this particular state of affairs must (via deductive force) also be true.* The later part, the very precise state of affairs, is sometimes called a "prediction." Putting it shortly, making a hypothetical claim is deducing, very precisely, a prediction, or to put it differently, deducing what the world must be like if the model were true. Let us use an example to illustrate and better understand these subtle points.

5.3.5.1 Experimenting With The Sun

Albert Einstein published his general relativity paper in 1915. Einstein's theory explained that mass deforms space-time pretty much (albeit in a multidimensional fashion) in the way that you can deform an inflated balloon by poking your finger into it. Of course, Einstein did not just say this, he mapped out precisely (that is, in a mathematical -- i.e. deductive -- way) the manner in which the modeled causes in such a system should ideally articulate. So, one of the conclusions that Einstein derived was that gravity is in fact the warping of space-time. Since Einstein's was a scientific model that meant that anybody using IDA would now be capable of producing hypotheses to test it. That is, out of his equations, anyone could deduce statements of the *if–then* form (conditional statements) that predicted what *should* the world be like if Einstein's models were correct. So, about 5 years later, in 1920, two British astronomers: Arthur Eddington and Frank Dyson, derived the following hypothesis from Einstein's models.[32] They argued that, if general relativity is the case, then a huge mass, like that of the sun, should warp space-time in such a way that its results should be visible under the right circumstances. The space-time (massive) dimple that the sun causes, they continued, should act as a distorting lens for any ray of light that hits our eyes after passing by it. For instance, this should be the case for any ray of light that leaves a distant star in a path that grazes the sun before hitting your eyes. Please do not try this experiment at home! Of course, you should not look at the sun directly, but even if you were silly enough to do it, you would burn your retinas for nothing. During the day, the light of any star that follows this trajectory

[32] An accessible account of this piece of science's history could be found at the following University of California, Berkeley's website: http://undsci.berkeley.edu/article/0_0_0/fair_tests_04.

would be overwhelmed by the light of our own star (the sun) and thus be invisible. This is why you do not see the stars during the day. But what if the moon were to be completely covering the sun, as during a full eclipse? That is precisely the circumstances that Eddington and Dyson used to make their prediction. They deduced two very telling measurements. Measurement one, to be taken during the full eclipse, would measure the precise location of a star while its light grazes the sun before hitting the instruments. Measurement two, during the night, would take the same reading but when the starlight is *not* grazing the sun before hitting the instruments. The two astronomers deduced that, if Einstein's model fits reality, then the two measurements should be different. But not different in any vague way, but different in *precisely the way that their calculations predicted*. We can summarize their argument as follows (where **EM** stands for: Einstein's model; and **DePo** for: mathematically deduced position given EM and the particular mass of our sun).

P1: If EM => DePo
P2: DePo
____
EM

As you now know, concluding EM this way is *not* deductively valid. As we have seen, experimental results *cannot prove* the hypotheses that generate them. Nevertheless, we also saw that scientists (along with the rest of us) consider a confirmation event very seriously. Now, how *objectively* this taking-seriously is, is the million-dollar question. As we have also seen, the answer to this question is not decided by how much the scientist likes the confirmation event, but on how much combined IDA support stands behind it. Now, back to the eclipse.

The 1919 eclipse was going to be visible from both sides of the Atlantic. Hoping to avoid bad weather spoiling the measurements, Eddington and Dyson picked *two* best spots, one on each side of the ocean, and planned two expeditions to observe the predicted results (see figure below). One team (lead by Eddington himself) went to the island of Principe (near Africa). The other team was sent to Sobral, Brazil. As the eclipse time neared, and in spite of Eddington and Dyson's best plans, the cloud cover (in both spots!) was thick. But right in time for the measurements, the clouds broke (in both spots!) and the teams of scientists got their *data*.[33] It matched

[33] A definition of this important term is coming immediately below.

their precise predictions, and at least judged by the media coverage at the time, this became one of the most sensational scientific experiments ever.

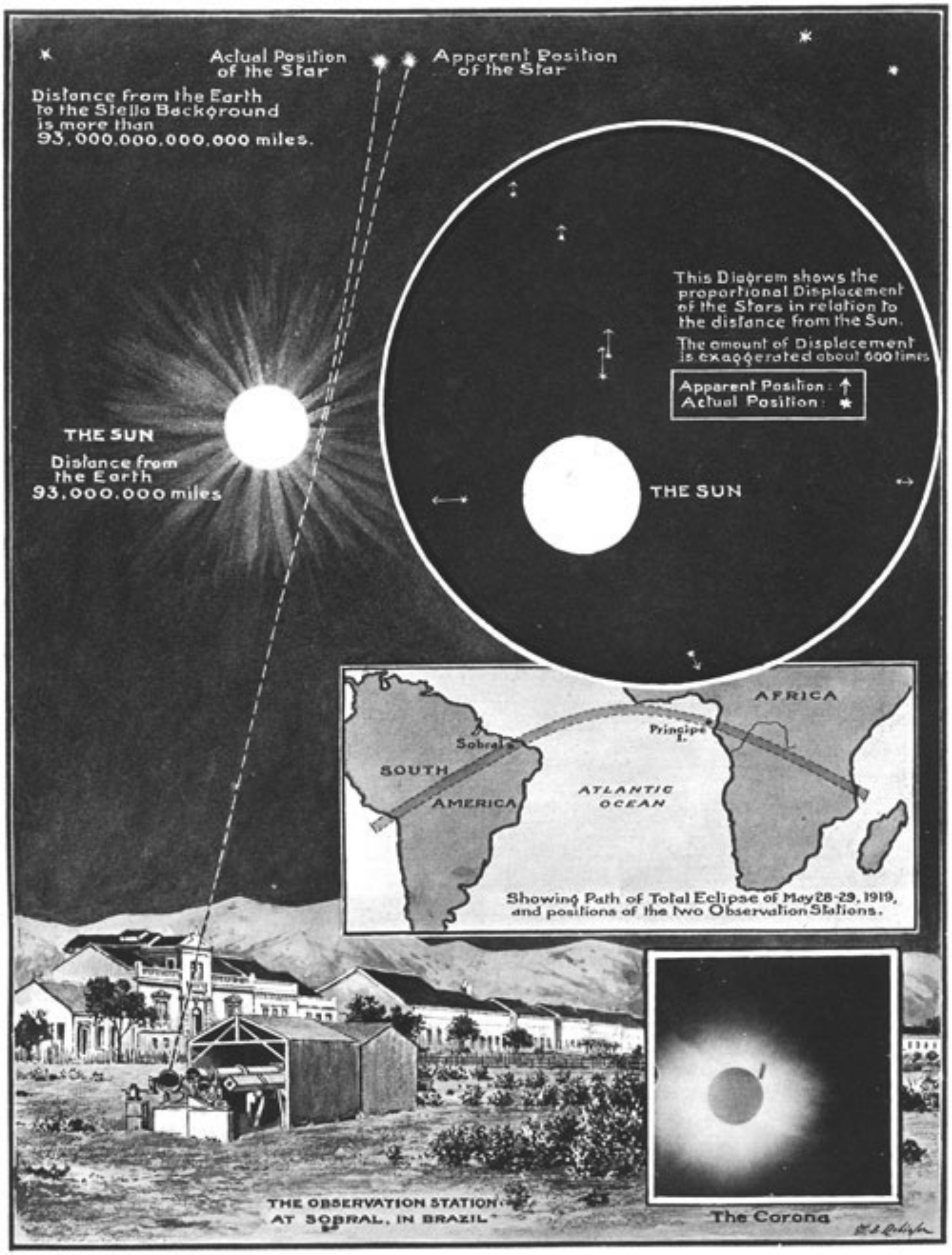


Figure 1. Above is the explanatory diagram that appeared in the London Illustrated News, on November 22, 1919. The caption read: ‘The results obtained by the British expeditions to observe

the total eclipse of the sun last May verified Professor Einstein's theory that light is subject to gravitation. Writing in our issue of November 15 [1919], Dr. A.C. Crommelin, one of the British observers, said: "The eclipse was specially favourable for the purpose, there being no fewer than twelve fairly bright stars near the limb of the sun. The process of observation consisted in taking photographs of these stars during totality, and comparing them with other plates of the same region taken when the sun was not in the neighbourhood. Then if the starlight is bent by the sun's attraction, the stars on the eclipse plates would seem to be pushed outward compared with those on the other plates…. The second Sobral camera and the one used at Principe agree in supporting Einstein's theory…. It is of profound philosophical interest. Straight lines in Einstein's space cannot exist; they are parts of gigantic curves.'" From *Illustrated London News*, November 22, 1919.

5.3.5.2 Experiments and Data

There are two ideas that we have been using quite a bit here, albeit in an unexamined way (which is always a bad idea!): *data* and *experiment.* We are now ready to unpack some important subtleties in these two terms, so let us start with *data*. Data, so to speak, does not "occur in the wild:" for data to exist there must be, first, models and hypotheses to generate it. It is important to realize that, without Einstein's modeling, Eddington and Dyson would have had nothing to measure when they looked up at the eclipsed moon. It is *only because* they had a set of precisely mapped causes (Einstein models), that they were able to precisely deduce a set of predictions. And it was only then that they were in a position to know two important things: what to expect, and what to measure. The term *data* refers to the latter: these are the actual, real world measurements that, thanks to the models, hypotheses, and predictions that generated them, the researchers can go out and obtain (or not). Thus, if the data measured or obtained is just as predicted, the hypothesis is considered supported or confirmed, and if it is not, it considered refuted. In Eddington and Dyson's example, if the data (the parameters in the actual world that the models and hypothesis sent us to measure) had shown no star displacement, or a displacement with a value other than the one deduced, then their hypothesis would have been refuted.[34]

[34] This illustrates, once again, how silly the idea of a refutation sinking an entire scientific theory is. If Eddington and Dyson had obtained the wrong displacement data, then a gargantuan piece of detective work would have been set in motion. The researchers would have started by trying, first, to figure out which was the weakest link in the complex set of assumptions, regarding the complex set of models that was used to predict the data they expected to observe. "Refuting" Einstein's Relativity theory from such a (imaginary) failed experiment would have been as arbitrary and premature as to refute instead (and/or also?) the theory

Now we should be ready to unpack the term *experiment*. An *experiment* is a setup that de-clutters the world in such a way as to (all other things being equal) infallibly bring a hypothetical prediction to actually happen. If it does, then the hypothesis is confirmed, if it does not, hen it is refuted. So, form the models, we deductively derive hypotheses that predict a particular set of hypothetical data, and experiments are the actual setups we devise (or are lucky enough to find naturally out there) where we can observe the predicted data play itself out, or not.

Experiments tend to be "unnatural," in the sense that such convenient setups are not easy to find and must usually be constructed. Obviously, causes out there work in conjunction. So, to isolate causes in order to study them, a scientist must intervene with the world, creating a situation in which unwanted causes and variables are barred from entering the picture and the target causes can produce the predicted data unobstructed. This would be the typical case of a laboratory setting: where the scientist constructs a physical setup in which machines and/or other props isolate the variable/s predicted by the hypothesis to vary in this or that particular way, and excludes (as much as possible) all the others.

For experiments of the natural kind, scientists take a non-interventional approach. In this case, she/he looks (through space and through time) until he/she finds a natural situation that (just like their artificial counterparts) isolates the variable/s under study. The later is the case of Eddington and Dyson's experiment. Natural experiments are the smart way to go when the variables under study are not easy to manipulate (like, for instance, a massive object such as the sun). Eddington and Dyson, so to speak, waited until a natural occurrence, the full eclipse of 1920, would alter the states of affairs just so that the effects of the target variables (the huge mass of the sun, the light of the stars, etc.) were de-clutered out of the otherwise noisy background.

5.3.5.3 No Irrefutable Hypotheses Allowed

We are now ready to extract several take-home messages from this famous example. First, not every conditional statement with a state of affairs at its right end (the prediction or *factual implication* end) can work as a scientific

---

that our sun has mass, or that we orbit it! In short, a refutation does not tell us that the whole set of models that produced it is wrong, but that *something* in this set went wrong.

hypothesis. If your factual implication is too vague, then your prediction is worthless. Let me show you this with a middle-of-the-road astrological slight of hand. If I am clairvoyant, then I should be able to predict the world, so I say: "tomorrow will be quite a day!" It should be obvious that *any* state of affairs will *confirm* this hypothesis (that I am clairvoyant). Or, to put it slightly differently, the clairvoyant hypothesis *cannot possibly* be refuted because nothing could happen that it would be contrary to it! Eddington and Dyson's hypothesis, on the other hand, was scientific because its prediction was not only about a state of affairs, but it was also precisely specified.

Now, you are in a position to understand why scientists thrive to make their models mathematically precise. A mathematically precise model should allow a scientist to derive precise predictions. That way, when the time comes to collect the data, there is no room for ambiguities: the hypotheses can be unquestionably corroborated, or refuted.

5.3.5.4 Not Every Model is a Scientific Model

Another message to take home is that not every cause can be modeled scientifically. Just like with the case of un-testable hypotheses, models that cannot produce scientific hypothesis cannot be assessed through IDA and thus cannot be considered scientific. For a causal map to be a scientific model it has to be able to mathematically/deductively derive, first, a healthy number of hypotheses, and second (from the latter), precise factual implications. Not every model can do this and this is the problem that typically plagues most pseudo-scientific attempts and bars them from being taken seriously.

Think, for instance, about the case of Creationism, or its emperor's-new-clothes version: Intelligent Design.[35] Creationism proposes a divine, a-natural creature (God) as the cause of all things. Clearly, nobody (not even God, I guess!) could make a deductive/mathematical model of such a divine entity and hence nobody can get the scientific engine started. Without a model, we cannot assess the support form above, or from below. Intelligent Design, even when its proponents claim it to be non-religious and different form

---

[35] A comprehensive (but of course ultimately hopeless) defense for Creationism and Intelligent Design as science can be found at: http://www.discovery.org/. I recommend you practice what you have learnt here by applying your new tools to dismantle the arguments there.

Creationism, suffers from the same problem. As long as the intelligent designer cannot be modeled scientifically, its proponents cannot claim to be able to generate any scientific support whatsoever (nor can they claim that their speculation is refutable in principle). "Science" is not a name that can be applied to just anything; it is a distinctive set of procedures that are not applicable to all dimensions of the human experience.

5.4 All Together Now: The Review of Our Peers

I hope that, at this point, you can bring back to mind our discussion of *plausibility* (Chapter 4). As you remember, determining the *best explanation* in an abductive move required one to account for two angles. First, the candidate cause had to be able to account for all the premises in question. And second, the candidate cause and this accounting had to articulate, or fit with all established background knowledge. Now, I hope, you are in a better position to understand what this plausibility check really entails, and where does this repository of background knowledge really lie.

As we have seen in this chapter, scientific thinking is simply IDA thinking with many extra layers of rigor and sophistication added to it. The accumulated results of scientific thinking, consequently, must be understood as that set made up of our most established facts, or what I have been calling here *background knowledge*. Now, one may ask, how can any human being perform such an immensely complex and interminable-sounding job?[36] The short answer is that your intuition is correct: no single human being can do it. The precise and complex job of checking the fit of new knowledge candidates against this immense background of accumulated scientific knowledge is performed by a vast number of experts, organized by area of expertise, and known as the *peer review* system.

---

[36] Just to give you an idea, a recent report -- Ware, M & Mabe M (2009) The STM report: An overview of scientific and scholarly journals publishing, Report for the International STM Association www.stm-assoc.org/2009_10_13_MWC_STM_Report.pdf -- estimates that there are about 26,000 scholarly journals that jointly put out about 1.5 million articles per year. Both the number of journals and of articles has been steadily increasing each year.

5.4.1 *Peer Plausibility*

Scientific societies have been around for a long time. For instance, one of the first modern societies, the "The Royal Society of London for the Improvement of Natural Knowledge," was founded in the year 1660, and it is still alive and very well today.[37] The AAAS (American Association for the Advancement of Science) was founded in 1848, publishes the prestigious *Science* magazine, and with more than 140,000 members, it is the largest scientific society in the world.[38] Scientific societies group scientists into sets by their disciplinary expertise. These disciplinary peers screen each other's results (using IDA exclusively) in several ways. There are special gatherings (conferences, symposia, or workshops in the field) where scientists submit for acceptance and then presentation and defense. But, most importantly, scientists screen each other's claims by submitting them for publication and, if merited, published. Scientists refer to these published writings with the misleadingly humble name of *papers*. Papers appear in magazines usually called *journals,* which can be highly specialized (like the *Journal of Exotic Pet Medicine*) or broad (like the above mentioned *Science*; one of the journals where the most important findings, in any field, tend to appear). Publishing in journals is not easy (specially the most prestigious ones). First, papers must be framed and articulated within a very complex network of established theories and results (both within and outside the discipline). Papers have to be written in such a way that highlights how anybody, using IDA alone and replicating precisely the same procedures, should obtain precisely the same results. And these filters will only take you up to the submission-ready state, the real hardship starts with the IDA screening by your best-established peers. After submittal and initial screening, the editor of the journal selects an anonymous group of most appropriate top-notch peers. This group of appropriate peer reviewers scrutinizes the submitted paper to determine (still exclusively on IDA grounds) whether or not the paper is offering something new that follows strictly and exclusively from the IDA-established background knowledge within and outside the field in question. But even if the reviewers agree and the paper gets published, the heaviest scrutiny has just gotten started. Once a paper is published, it is now open to the scrutiny of any other peer scientist (and, for that matter, the public at large). If anybody spots an IDA problem with the paper, then

---

[37] You can visit their website at: http://royalsociety.org/.

[38] You can visit their website at: http://www.aaas.org/.

they themselves write it up, submit it for publication, and the whole cycle starts all over again. Thus the cumulative formation of scientific facts gets in motion.

5.4.2 Scientific Theories are Super Facts

As you can see, the peer review process is a machine-like superstructure where scientists, working individually, or in collaboration, are the cogs, or elemental units. Each individual cog might not be objective, but the structure of the peer-review machine is such that it maximizes the IDA screening of the products that pass through it. Scientific theories are the superstructures of phenomena and their causes that congeal out of it. To put it a bit more precisely, the peer-review system takes in hypothetical puzzle pieces, or clusters, regarding hypothetical correlations, effects, and causes in nature (that scientists submit). Clusters loose or add new pieces to them, and might occasionally disappear altogether. Some clusters prove to be sturdy and stable, and begin to congeal into a causal map of nature. The latter are our scientific theories.

One must bear in mind that this system is made out of humble, fallible humans. This realization is of the uttermost importance to appreciate how successful science is as a fact-finder -- specially when compared to other putative fact-finding types of human endeavors (like politics; or religion). The peer review process does not require its practitioners to be unusually virtuous, or inhumanly objective. In fact, we know that they are not. Scientists should be expected to be just as objective as your average human being. But we could conceive the efforts of individual scientists as the fibers and threads that through the peer-review process get selected (or not), cleaned up, weaved, revised, and re-weaved into ever more comprehensive and interconnected pieces (the scientific theories). Thus, objectivity emerges as the combinatorial, communal effect of more-or-less objective peers checking the IDA of each other.

There are at least two rather surprising implications that come out of our analysis. First, since we (justifiably so) consider facts more or less established depending on the amount of IDA behind them, then scientific theories must be considered as *super facts* -- specially when compared to such mundane facts like table, chairs, or the piece of paper or screen that you're reading this essay on. Second, you might ask the very good question of where do scientific theories reside. Facts about tables, chairs, and screens reside in your head. But where do scientific theories (super facts) lay? Is it in the individual scientists' heads, or in the communal, hive mind (to call it somehow) of the disciplinary and interdisciplinary groups of peers, with their journals, and their meetings? Let us figure these two excellent questions out.

5.4.2.1 Super Facts

I argue that the peer review system produces *super facts*; or *facts on steroids*. There are two main reasons behind me making such a strange (but justified) claim. One, as explained above, publishing is all about IDA and little to nothing to do with the individual idiosyncrasies of the particular scientist doing the publishing. Two, even when particular scientists let personal idiosyncrasies play an undue part in their research, exposing the research to the peer review system will, sooner or later, allow IDA to weed out the transgression.[39] Having a diffused net of IDA nodes (the peers) monitoring the work of each other (instead of central monitoring node) ensures that, even with an expected constant influx of human subjectivity entering the system (scientists, after all, are human beings), objectivity should rule in the long run.

Now, that settles the "fact" label for scientific theories. As we have established earlier, a *fact* is a conclusion that comes from most excellent IDA moves. We have also shown that the reliability, or how seriously we take the fact in question, should be directly proportional to the amount of IDA backing it. Finally, we have also demonstrated that the peer review system is designed to maximize the operation of these critical thinking tools – the only ones available to us humans. One should most definitely take the existence of the tables and chairs seriously when one has tons of IDA supporting the fact of their existence (but remember also that all the IDA in the world will *never prove* their existence – but, as we have also seen, *proof* is not only unattainable but also greatly overrated). So, tables, chairs, and the remaining furnishings of the world around us are pretty solid facts because they are backed by pretty solid evidence.

But what about scientific theories; what about electrons, neutron stars, black holes, or the force of natural selection and plate tectonics? Granted, their existence feels quite a bit more questionable than that of tables and

---

[39] This is demonstrated by the occasional withdrawal of published research from learned societies' journals after scientific malpractice, or other IDA violations are discovered by the peers.

chairs. Unlike their humbler cousins (the tables and the chairs), the causes in most scientific theories cannot be touched, seen, smelled, or tasted (please don't even try to do anything of the sort with black holes or neutron stars!). Naturally (no pun intended), our ancestors were selected to take most seriously the stuff that they could see, touch, smell, and/or taste (probably in that order of importance). But we, their descendants, trying to think outside the constraints and biases of that natural box, should see that this historical baggage is little more than a practical solution designed by natural selection – not a piece of careful, IDA thinking. Our favoring of the touchy-feely type of facts results from the local context of our ancestors and the process that shaped their cognitive capacities. It is a historical fact that our ancestors' most reproductively savvy decisions were consistently (and obviously most directly affected) by mundane causes such as potential mates, leopards, friends, and foes. What these causes had in common was that they could all be seen, touched, smelled, and or tasted thanks to the sensory systems available to them. Quite obviously, quantum effects, the distribution of dark matter, or the existence of the Higgs Bosoms could be clearly disregarded when it came to questions of Earth's life forms reproductive/survival success – or lack thereof. In fact, while playing the reproduction/survival game, one could argue that too much of the latter kind of highly-theoretical thinking must have consistently been a poor gaming strategy, and that in fact it must have been easily bested by more street-wise, or down-to-earth thinking.

Putting it shortly, our cognitive skills were made to be highly sensitive to the mundane-kind of causes, and not very apt to be terribly moved by those closer to the highly theoretical kind.[40] But this historical baggage should not be confused with some sort of intellectual justification for such a bias. As we were reminded above, a fact's reliability should result exclusively from the amount of IDA inferences behind it. So, if the peer review system incarnates the most scrupulous exercising of IDA thinking possible to us, then peer review systems produce the strongest facts available to us. Thus, and I know this should sound counterintuitive to you (this is natural selection's bias at work!), black holes, the force of natural

---

[40] A very perceptive question, at this point, would be: "But what about our beliefs in the supernatural, and what about religion?" You might enjoy reading philosopher Daniel Dennett's excellent book "Breaking The Spell," for an excellent theory of why this particular kind of non-evidence-based thinking leap-frogged the natural selection concerns outlined above.

selection, or any established scientific theory has *more* evidential support behind it, than the tables and chairs laying in front of us. Consequently, if the latter are (rightly so) considered solid facts, then the first should also be rightly considered to be *super* solid facts.

5.4.2.2 Where Scientific Theories Lay

The second surprising result of our analysis that I would like to bring your attention to regards the question of where are scientific theories housed. This is an important question because its answer should determine whom should you listen to. Or, putting it slightly differently, it will allow us to know who keeps the super facts so we can go and ask her what they are.

Remember that established scientific theories are very complicated families of models (sometimes, with more than one competing model applying to the same thing!). And families (theories) connect with other families, far and wide. How far and wide theories connect is hard to envision, specially, at the same time one tries to bear in mind that this connection takes the form of a very strict relationship. Theories delimit and constrain each other. Theories in physics, for instance, tell all other theories "above" them (like those in economics, psychology, or cosmology) what can and cannot happen. Similarly, (neuro) biological theories frame the possibilities of (for instance) psychology, education, and or sociology. Amazing visualizations of this interconnectedness can be glimpsed in the work of Richard Klavans and Kevin W. Boyack,[41] or Figure 2.

[41] For instance, see: Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60(3), 455-476. Preprint Published version

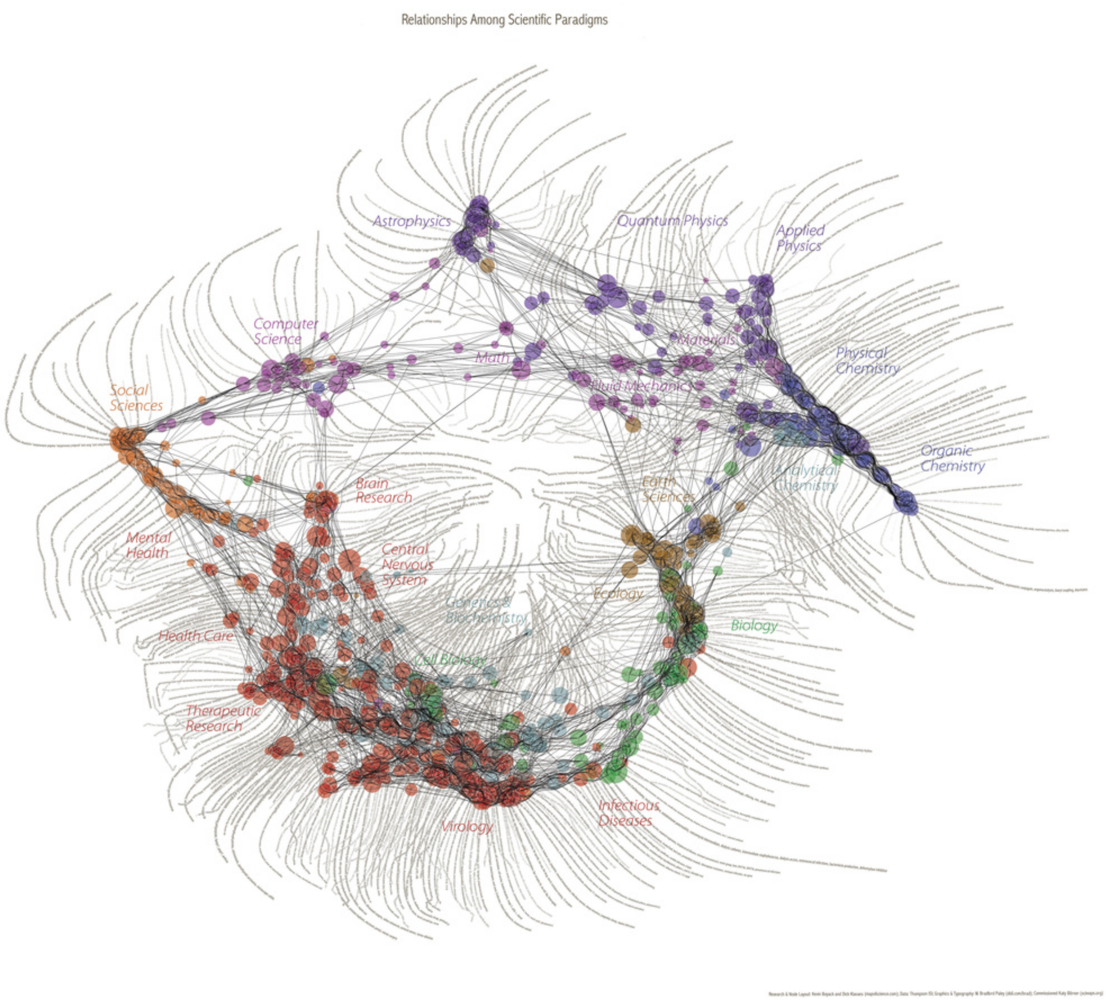


Figure 2. This map connects the theoretical backgrounds that different papers brought to bear while establishing their points. The website where it appears allows you to blow it up and scroll over the entire thing. It would be an excellent idea to go to this website where and see how disparate these theoretical bodies are. http://seedmagazine.com/content/print/scientific_method_relationships_among_scientific_paradigms/

Bearing this constraining interconnectedness in mind, it should be easy to see why a single human brain cannot possibly hold, let alone IDA-check, full-blown scientific theories. There would simply be not enough room for the necessary software and hardware to do so. Producing and conceiving scientific theories must be at least a multi-brain kind of endeavor. Indeed, scientific theories emerge from the peer review process and they are housed in the combined total of the publications that disciplinary and interdisciplinary scientific societies produce. So, in a very important sense, scientific theories do not live in any particular place (or head). They live in the consensus of peers, and are materialized and developed in their societies' meetings and publications. It is conceivable that, at a very early stage in the life of a given scientific theory, it could have been housed in a single individual head (think, for instance, of Charles Darwin's theory of evolution, or Newton's mechanics). But add the actions of peers to the mix, coming up with new IDA ideas, criticisms, and tests, subject their products to the peer-review process, and sooner than later the emerging theories are too big for any given single human head to hold.

This separation between the scientist's head and the scientific societies' "heads" should help you make sense of many of todays unfortunate science and technology "debates." For instance, it can explain why the media (operating at its very worst) is able to perpetuate myths of scientific controversy when it comes to some problems that happen to be politically charged. Think, for instance, about the "debates" between evolution and intelligent design, or the denial of the human role in global warming. Unscrupulous (or maybe simply confused) journalists can always find some outlier scientist that does not accept this or that peer-reviewed, established theory. This should be surprising only in the sense that the scientist/s in question has/have not yet realized that theory acceptance is the business of the community of peers, and not a matter of personal choice. More importantly though, considering the existence of confused scientists (and journalists, and publics) as an indicator of scientific controversy (that is: a real controversy at the the peers' level) simply denotes a fundamental misunderstanding about what scientific theories are and what their establishment is all about.

5.5 Observing the World through Theoretical Eyes

Up to this point, you could still think that all this talk about the nature of scientific theories is probably interesting (I hope!) but only in a limited, purely academic way. After all, you might think, scientific theories deliver a map of reality that has little to do (if any) with the map our senses deliver. Furthermore, the latter map is the one that really fits our common sense and our daily experience. Thus, understanding the role of theories in general, and scientific theories in particular, you might conclude, does not have much import beyond the scholar's armchair. But forgive me in advance because I'll argue below that you are quite wrong. As it should be the case with every good academic problem -- and this one is a good one, understanding what theories are will produce many tangible and important implications. Most importantly, I will demonstrate that perceiving is impossible without theories: that in order to perceive the world one needs to deploy an array of theories. So profoundly so, that if I were to magically erase all theoretical devices from your head right now, you would be utterly unable to perceive, or experience anything. Consequently, if deploying theories is inseparable from the act of perceiving, questioning the quality of such theories is crucial if we are to bet our lives on the veracity of the perceptions. So, asking ourselves what makes a good, objective theory should now be understood as one of the most important questions we could possibly ask.

Let us start, first, with the difference between *observing* and *seeing*.

5.5.1 Observing versus Seeing

These two terms are usually used indistinguishably. But there is an important distinction hiding here. Even colloquially, the action of *observing* denotes a few more degrees of involvement than that of merely *seeing*. When, in plain English, we say that we observe something we mean to stress that we're putting some effort into the action. There is some wisdom in this colloquial distinction. Let me unpack it with the help of an analogy.

Imagine a lens, any lens, like a magnifying glass, for instance. Let us add it to a box, so that the lens is inserted in a dedicated whole on one of the box's sides. The light coming through the lens would hit the opposite wall of the box, right? Now, what I would like you to concentrate on is the question of what would it take to turn the lens-plus-box in our example into a, let us say, night-vision camera, or better yet, into a motion detecting device. The rather simple answer is that you would need to add the right kind of software and hardware to our humble box and lens combo. The important thing to notice here is that there are two very different jobs being performed here. The first job is the one accomplished by the lens and the box, and is that of bringing the light through the lens and onto the opposite wall of the box. This job I will call the job of *seeing*. The second job is that performed by the software and hardware. This is the job of analyzing the inputs above, and I will call it the job of *observing*.

Thus, observing involves *processing*, or the transformation of the input into something else. More precisely, it requires taking in the input (light, in our example), subjecting it to some kind translational procedures, and outputting some sets of results. Think, for example, about the smile-recognition capacities in many of our digital cameras. The lenses *see* in the sense that they take the input light and transfer it to other hardware and software that digest it (to call it somehow) via a series of procedures that transform it into a different output. In the case of the smile-recognition cameras, a computer chip translates the image borne by the lenses into a particular string of zeros and ones that gets to be manipulated to produce the output or *observation* of a smile – or lack thereof. Now, let us see what can this framework of analysis tell us about our own heads and the distinction between seeing and observing, and at least as importantly, what does all this have to do with scientific theories.

Think about our eyes and other sensory devices (your nose, your tongue, your skin, etc.) as the input-channeling devices in the previous examples as the elements that do the seeing.[42] and our neuronal networks as the systems performing the translational procedures that would crunch the sensory input into their outputs: our conceptions of the world. Just like with the examples of the lens on a box, the motion detector, or the smile-recognition camera, the same three key elements are identifiable: the input signal; the input signal processing or crunching procedures; and the output, or end result of the processing. With these distinctions on the table, we can now tackle effectively the very important question of what makes an observation *objective*.

Let us start at the opposite end of the previous question: what would it take to come up with a mistaken or *subjective* observation? Certainly your eyes (or other input channels) could malfunction. We deal with this kind of problem all the time when we prescribe corrective glasses. But I would like to consider a far more interesting, at least for our purposes here, set of circumstances. What if your eyes work perfectly well, but your crunching or translational formulae are all messed up. In that situation, you would be *seeing* right, but you would be processing your inputs wrong: then your output would be off, and you would be *observing* wrong. Fixing this problem would be a bit more difficult. Depending on the magnitude of the processing malfunction, fixing the output, making you observe the right thing, might require extensive re-education or, putting it a bit differently, major re-

---

[42] This is only partly correct. The distinction between seeing and observing, when it comes to your eye, for instance, is a bit more blurry. Your retina, in fact, is already doing some processing.

programming of the procedures that are crunching the inputs in the wrong way.

We are now in a position to be very precise about what it takes to make objective observations. When the input devices work properly, and when the input processing formulae are good, then we can say that the observations made are objective. But if either the input devices, or the crunching devices applied to the input feeds are bad, then we would say that we are observing some level of nonsense, or that the observations we make are *subjective*. So we can now meld our discussion about fact-finding with our understanding of the ubiquitous role of theories. The question of whether or not our observations of the world are objective, or whether or not we are getting the facts right, boils down to a question of whether or not we are observing the world with good input devices and our most excellent theories.

5.5.1.1 Theories & Objective Observation

Philosophers refer to the insights above when they stress that no observation is *a-theoretical;* or that every observation is *theory laden*.[43] As we outlined above, the term *theory* refers to the computational maneuvers that get you from input to output. Or, putting it slightly differently, we are calling *theories* those computational processes that convert a certain set of sensory inputs into an intelligible observation.

For instance, we conclude, or *observe* a flowerpot; a dog; the tundra; or the taiga, when we are presented with certain sets of input patterns and we crunched or processed them by certain sets of theoretical maneuvers. The first important thing to notice here is that many of the theories that are brought online to conclude any observation lay way below the limelight of our awareness.

The theoretical apparatus necessary to spot a cat might seem pretty obvious, and frankly, sound also pretty unimpressive. But then, you better think again. Think, for instance, that in order to spot any*thing* one must first, for instance, be able to spot *things,* in general. This might sound like a mundane task but programming (i.e. providing the correct *theoretical*

[43] Most treatments consider under the word "theory" both input devices and theoretical devices. Here, they were detangled for didactic purposes. Naturally, input devices can be also conceived as part of the theoretical processes.

devices) a computer with the code to isolate *things* out of a *background* has proven to be a daunting endeavor.[44] So, understanding the connections between observation and theory is not easy, but it is hard to overemphasize the far-reaching implications that this understanding has.

To help us detangle the interrelatedness between observation and theory, let me propose the following analogical aid: let us conceive theories as *sophisticated, science-fiction like goggles*. Unlike any old pair of protective goggles, imagine these theory-goggles as the kind people might wear in a futuristic film. These goggles would allow them to see, let us say, electromagnetic radiation (e.g. light, radio, or gamma rays) in ways that no previous primate would have ever dreamed of. But before we understand such futuristic devices, before downing our scientific-theories goggles on, we must first understand that we are already wearing a pair of very much not futuristic goggles. We are wearing our very own permanently implanted, evolutionarily made, standard-issue monkey goggles.

5.5.2 Monkey Goggles

We are monkeys, and like every other species in our lineage, we rely heavily on vision. We must see it to believe it (not smell it – like a dog might, or electro-chemically taste it – like a catfish might, or hear it – like a bat might). I suspect that this fact is at least partly responsible for the reluctance that so many people has when it comes to accepting the causal map of the world that scientific theories disclose. The world of deep causes that science (like the futuristic goggles above) allows us to glimpse is vastly unlike the one humans can naturally *see*. Furthermore, most of science's causes are simply "un-see-able" ("*unobservable*" is the technical term). That means not that humans do not usually see the causes that science can see, but that humans *cannot possibly* see them!

Some might feel that if the world that science discloses is very much unlike the one that we can see with our plain eyes, then there must be something wrong, somewhat deficient, or less certain about the scientific picture. But in fact, if one were to think about this carefully, one should realize that expecting a deep overlap between the world as disclosed by science, and the world as disclosed by our senses, would be quite presumptuous. Imagine the chances that some insignificant species, in some insignificant corner of

---

[44] Watch, for instance PBS' *Scientific American Frontiers*' "Robot Pals," at: http://www.pbs.org/saf/1510/video/watchonline.htm.

some insignificant backwaters galaxy, would evolve a set of sensory and theoretical devices that were capable of capturing the fundamental nature of reality! As we have seen, evolution shapes adaptations (e.g. sensory and theoretical devices) in a very myopic way. Firstly, natural selection can only tinker with whatever variation happens to be at hand (it cannot build functional apparatuses from scratch). Secondly, it can only do so within engineering goals that are restricted, *exclusively* to: (a) matters of reproduction (and survival, but only in relation to reproductive success), and (b) the locale (in terms of space and time) where the survival for reproductive success in question is happening.

With the above in mind then, one should actually be mightily surprised if there was a great deal of overlap between the two pictures. Of course we should expect some overlap when it comes to issues that were crucial to our ancestor's reproduction and survival. But, for the same very reason, we should expect the differences between the two conceptions grow farthest, the farthest we move away from these areas of myopic Natural Selection's concern. Imagine for a second an alternative reality. One were a process they call "science," a process that is purported to deliver an objective map of the world, discloses one that, at its most elemental level, works precisely as our locally evolved, reproduction/survival concerned, idiosyncratic sensory and theoretical apparatuses tell us. In such an alternative reality, you would have every reason to believe that this "scientific" map has not been derived via IDA, that it has been fabricated for some subjective reason. This miraculous match alone would indicate that whatever it is that they call "science" in this alternative reality, is more concerned with delivering what we *want* to see than what actually *is*.[45]

Thus (and I know this is counter intuitive), we should *not* expect the most profound maps of reality that science discloses to be human-cognition friendly. As you should now clearly see, scientific theories should be conceived as our best efforts at *removing* our human-selves from the picture. Not surprisingly then, when we succeed in doing so, what we get

[45] In fact, you do not have to travel to imaginary universes to find instances of what we have just imagined. Not surprisingly, most mythologies about origins place the mythmakers as the chosen ones, at the center of the universe (both in terms of geography, and in terms of relative importance) and possessing cognitive capacities that are not severely biased by their biology, but, on the contrary, are unlimited and capable of perceiving and conceiving the world precisely as it is!

is not quite monkey-friendly at all. This is where the idea of goggles comes in. To help us understand and celebrate this unfriendliness, let me propose a different way of conceiving our relationship to our permanently implanted, monkey goggles (our natural-selection shaped cognitive apparatuses), and our scientific theories. I will invite you next to think of scientific theories as intelligently-designed *observational* devices: as something like sophisticated goggles, or virtual reality helmets, that you can put on, to see the world (for the first time in our species' history!) *not* from behind monkish eyes.

5.5.3 Theories as Goggles

When one puts a pair of night-vision goggles on, one gets to observe the world at night better. That is the case because goggles are devices that were designed to crunch a certain input that our eyes are not good at capturing and to produce a certain output that our eyes are good at capturing. Good goggles can do this because they have the necessary software and hardware to gather the little light that hit their lenses; the ability to process it, and to translate it in a way that makes sense to us. Now, let us consider the following question. What makes a pair of night goggles a *good* pair of night goggles?

To answer this question, it is important to realize that goggles are nothing more and nothing less than the instantiation of a very complex set of IDA maneuvers. Some of them still bear signs of their original form; that is they are logical moves in the form of the goggles' software. But some others, the goggles' hardware bits, are physical contrivances that contain, in their nature and their deployment, the pieces of reasoning that conceived them: they are instantiations of IDA maneuvers. In short, goggles are actually complex sets of IDA moves (some still "soft," in the form of pieces of code, and some "hardened," as they have been translated into physical structures) that turn their input or non-visible elements (their "premises") into visible outputs (or, "conclusions"). Goggles, in this sense then, could be fruitfully conceived as *theories*; since their output is model that is computed out of their inputs. By the same reasoning, *scientific theories* could be fruitfully conceived as goggles. Unlike the latter though, scientific theories do more than just allow us to *see* better; they help us *understand* what would otherwise be not only invisible but also rather unintelligible.

It is at this point where the goggle analogy, with its stress on a *visual* conversion might become confusing. As we have discussed above, scientific theories do not help us *see* the world behind the phenomena. Instead (and better put) scientific theories allow us to *causally map* it. So, pushing the goggle analogy just a bit, I am proposing here we understand scientific theories as some sort of *hypermedia* goggles, or virtual-reality headgear. The added media channels here are not necessarily audio or visual: the scientific-theoretical mediatic experience delivers *understanding*. When you don on your pair of scientific-theories goggles, what you see is a map of the causes that are responsible for the phenomena of the world. In short, imagine scientific theories as one of those goggles, or even better, virtual-reality helmets that tend to appear in some sci-fi movies. Scientific theories are understanding-goggles/helmets through which we see the world as a map of the causes behind the regularities of nature. These goggles are made up of sophisticated pieces of IDA, some in the shape of software (e.g. formulae tying phenomena; causal models; data; and apparatuses' design) and some in the shape of hardware (the apparatuses in question: all the measuring and crunching devices that scientists use). When you look through them, we begin the journey to understand the world around us in an objective way.

Now, you might ask, why do I say "we begin," in the previous paragraph? I say it because scientific-theories goggles work as corrective vision devices: they correct the causal maps that our personal evolutionary processes shaped for our species. Scientific theories allow us to grasp causes with as little evolutionary baggage as possible. Probably for the first time in the history of life in this planet, we have theoretical devices that can take us as far away as humanly possible from the limitations, biases, and idiosyncrasies ingrained in our minds and cultures by our ancestral evolutionary journeys.

As we begun to argue above, you and I have been wearing theoretical goggles all of our lives. And so have our parents, our grand parents, and every one of our ancestors, going back as long as the world outside needed to be regarded and modeled to make sense out of it. As we have seen, natural selection, the designing mechanism of these cognitive devices, is beyond stupid: it has no mind at all. That means that it cannot possibly design following global quality standards such as objectivity. As we stressed, natural selection design standards are exclusively provincial. It can only design structures, or theoretical devices out of the local pool of available variants (where new, raw variation is random, or undirected), and given the local standards of reproductive and survival success. In other words, natural selection could have only designed a set of theoretical devices, that we can call *monkey goggles,* which allowed our ancestors to deal with their local reproductive-survival problems in such a way that maximized their reproductive output.

These are not good news. A wise and intelligent designer would have designed our set of theoretical goggles following global standards of objectivity. But natural selection can only build goggles that are hopelessly biased and functionally restricted. Its designs are limited by such things as our ancestors' historical baggage; the materials and developmental pathways available; the luck of the hereditary-variation draw; and natural selection's very myopic set of quality standards: local problem solving -- measured exclusively in terms of reproductive output currency.

So, when it comes to natural selection, IDA is a slave of the prime directives of reproduction and survival. Monkey goggles should be expected to provide us with the kind of conceptual grasp of the world that best suited the reproduction and survival of our ancestors. Or, putting it slightly differently, the degree of fidelity, resolution, global comprehension, and/or enlightenment that these cognitive devices have is that which was required for their reproduction and survival. And, of course, you must know that pure IDA will not necessarily help you when it comes to making reproductive decisions.

5.5.4 Human Goggles

Once we did not understand the process that designed the monkey goggles – the ones that are permanently attached to our heads. We once considered our view of the world unbiased; our permanent goggles objective windows into the universe. But we should now know better. We cannot expect monkey goggles to be objective IDA devices. If we want the latter, we need to design them intelligently. As we have seen, scientific theories are precisely that kind of IDA goggles.

Are the IDA goggles of science entirely objective? As we have shown, such god-like objectivity is actually inconceivable. It makes no sense to imagine that, given our humble origins, our brains (or any natural-selection made brain, for that matter) could be able to even conceive what pure objectivity could possibly be. What we can hope though, is to be able to dull the edge of our monkeys' bias; to realize our very clear limitations, and aim for a causally competent understanding of whatever part of the universe our primate brains can grasp. We can call these enhanced monkey goggles "*human goggles*" (that is: monkey goggles with the scientific theories upgrade installed). Monkey goggles showed us a world distorted by our ancestral core design directives of reproduction and survival. Human goggles hold the promise to open up a brave new world of intelligent directives. Which are those -- you might ask? I am not sure, what is clear though is that we will have to figure them out by lifting ourselves by our own bootstraps. What I have just done though is my best to explain what are our best (and sadly our only) tools to discover them.

5.6 To Take Home

Remember way back, our example of the tiger gnawing at the door? With the help of the Glieseans we learned some surprising aspects about facts. First, we discovered two important, yet not very obvious things: that facts are the conclusions of arguments; and that, without facts, actions are not possible. We also realized (certainly to our dismay) that definitive certainty about facts is forever beyond our reach. Luckily, we also discovered that this ultimate failure should not paralyze us. If it does, then it means that we have been thinking about knowledge from the wrong perspective. Certainty is greatly overrated. Facts are crucial, but we cannot take them or leave them because we know for certain that they are true or false. We take them deadly seriously because of the process that produced them: facts are accepted as such when they result from an objective fact-finding process. In other words, we figured out that facts should be accepted or not depending the goodness of the reasoning that backs them up. Putting it even more shortly: facts are what IDA says. With this discovery under our belts, we moved on to figure out IDA thinking in order to maximize it. It was during this particular part of our journey, that we realized that science is nothing more, nor less than such IDA-thinking maximization. Scientific theories are the result of super IDA thinking, and consequently, scientific theories *produce* super facts.

The overall and final conclusion of this essay is both humbling and of great import. Now that we understand how you and I use our everyday-IDA thinking to come up with facts, and how peer review systems use their enhanced-IDA thinking to come up with scientific theories (or super facts), we have to conclude (via IDA, of course, and regardless of what you and I believe) that, if we live and die by our everyday-IDA facts, then super-IDA facts must be taken *at least as seriously*. Thus, understanding that the degree of IDA thinking behind a conclusion (and not its ultimate truth value) is what determines the degree of responsibility for its acceptance, it is our responsibility to take scientific theories (our only super facts) super seriously.

6 SOME IDEAS FOR FURTHER READINGS

- For more on critical thinking you can read Irving L. Copi & Carl Cohen's "Introduction to Logic" 13th ed. (2008) Prentice Hall.
- For more on the relationship between critical coomon sense and scientific thinking, you could read W. V. Quine & J. S. Ullian wonderful little book "The Web of Belief" 2nd ed. (1978) McGraw-Hill Humanities.
- For more on the nature and scope of science you might want to look at Ronald N. Giere's "Understanding Scientific Reasoning" 5th ed. (2005) Wadsworth Publishing.
- For more on evolution and the operation of natural selection, you could read Charles Darwin "The Origin of Species" (along with everything Darwin, and online, at http://darwin-online.org.uk/). Or you can read any of Richard Dawkins' wonderful books on the subject (like "The Selfish Gene" 30th Anniversary ed. (2006) Oxford University Press.
- For more on the idea of scientific thinking and its progress you can go to: http://plato.stanford.edu/entries/scientific-progress/.
- For an excellent historical introduction to the philosophy of science you should check out: John Losee's *Philosophy of Science* 4th ed. (2001) Oxford University Press.