A Response to Richard Carrier's Review of *C.S. Lewis's Dangerous Idea* (2007)

Darek Barefoot

Before I comment on Richard Carrier's review of Victor Reppert's recent book[1], *C.S. Lewis's Dangerous Idea: A Defense of the Argument from Reason* (InterVarsity Press: 2003), it is only fair to note that while Carrier is an atheist, I am a Christian. So it is no surprise that we start out on opposite sides of the book's central issue: a naturalistic worldview's compatibility or incompatibility with the existence of our faculty of reason. That doesn't mean that I find all of Reppert's arguments to be persuasive and all of Carrier's criticisms to be off target. Like Carrier, I find Reppert's arguments against the adaptive advantage of a large brain unconvincing. And I concede that different thoughts probably entail distinguishable physical events in the brain that are detectable in principle, even when they do not make an overt difference to behavior. But these are quibbles. I believe that the core of the argument from reason (AfR) is sound and that Reppert's book is a landmark contribution to the subject.

In evaluating the book, we should note that Reppert intends *C.S. Lewis's Dangerous Idea* (henceforth *CSLDI*) to be a sketch and summary of the argument rather than a developed formal treatment. In simple terms, the argument from reason turns on the observation that all events occurring in the natural world--from leaves rustling in the wind, to water flowing downhill, to gears turning in a watch--are in and of themselves "nonpurposive and nonintentional," that is, blindly mechanical. No matter how many events we contemplate, or how complex their interrelationships, this "blindly mechanical" description must apply in order to hold them all within the category of nature, and to see them as subject to the explanations of hard science. By contrast, the same description cannot apply to events in our minds, at least not when they form a logical train of thought. How can we explain the enlightened and goal-directed nature of logical thinking unless we expand our definition of what is real beyond the natural, physical world to include some "thoughtful" bedrock of reality, such as God?

Although I find many of Carrier's objections to Reppert's book (and to the AfR in general) to be confused and unsuccessful, I will begin with what I believe to be the central issues--intentionality, representation, and mental causation. I will then take up the secondary topics of epiphenomenalism and volition, causal closure and physics, the laws of logic, and the problem of induction.

Intentionality and Representation

Intentionality refers to the quality of "aboutness" possessed by our thoughts. In *CSLDI*, Victor Reppert points out that thoughts can be about things, but we do not speak about one bit of matter being "about" another bit. How can this strange quality be explained naturalistically? Carrier attempts to answer this challenge, but he invariably falls back on the very concept he is trying to explain. He stumbles into this trap again and again, despite Reppert's specific warning about it in the book (*CSLDI*, 119). For example, Carrier specifies that,

> Cognitive science has established that the brain is a computer that constructs and runs virtual models. All conscious states of mind consist of or connect with one or more virtual models. The relation these virtual models have to the world is that of corresponding or not corresponding to actual systems in the world.

He later elaborates:

> "This bit of matter is true about that bit of matter" literally translates as "This system contains a pattern corresponding to a pattern in that system, in such a way that computations performed on this system are believed to match and predict the behavior of that system."

To begin with, Carrier overreaches when he claims that cognitive science has proven that the brain *is* a computer. The current consensus, called functionalism, holds that the brain is functionally similar to a computer in certain respects. The theory that the brain simply is a computer--computationalism--is more controversial. In any case, cognitive science has not given a consensus response to critics of functionalism such as John Searle and David Chalmers, whose objections apply doubly to computationalism. So it is far from settled scientific opinion that the brain is nothing more or less than a computer.

Turning now to Carrier's explanation of intentionality, we must assume that the "virtual models" and "patterns" he refers to consist of series of synaptic discharges in the brain, since these discharges are the physical events that scientific instruments detect. But what does it mean in physical terms to say that such a series "corresponds" to an "actual system"? This is what Carrier needs to tell us. Let's draw an example from things outside of the brain that seem to have intentionality or aboutness--namely, sentences. A sentence can be about something, but it is difficult to peg this quality to a physical property. If a sentence is audibly spoken it can be loud or soft, or pitched high or low, without a change of meaning. The intentionality cannot be in the specific sounds, either, because the sentence can occur in a number of human languages and even the electronic beeps of Morse code. If the sentence is written, it can take the form of ink on paper, marks in clay, or luminescent letters on a computer monitor. The shapes of the letters are a matter of historical accident and could easily be different. The sentence can be encoded as magnetic stripes and as fluctuations in electrical current or electromagnetic waves.

There seem to be no physical limits to the way a sentence might be represented, provided that the representation is agreed upon by human minds, and therefore no physical property in which the sentence's meaning and intentionality reside. The physical form that the sentence takes can be arbitrary because it represents something other than itself--a thought or idea or proposition. A sentence saying that the Earth revolves around the Sun does indeed "correspond" to an actual state of affairs, but that fact alone does not demonstrate that the correspondence is a physical property of the sentence. This observation also applies to the correspondence of "patterns" or "virtual models" in the brain to the states of affairs that they are about. Contrary to Carrier, simply referring to "aboutness" in terms of correspondence does *not* demonstrate it to be physical[2]. Notice how another of Carrier's statements repeats the same mistake:

> The fact that one thought is about another thought (or thing) reduces to this (summarizing what I have argued several times above already): (a) there is a physical pattern in our brain of synaptic connections physically binding together every datum about the object of thought.

Carrier even uses the phrase "every datum *about* the object of thought" [emphasis mine], perhaps forgetting that "about" is what he is trying to define. If Carrier had shown us how some physical property of brain chemistry causes data in the brain to be about an object of thought, then a physical pattern might be assumed to bind the data together. Clearly, a thought's being "about" something entails a correspondence, match, or connection. But can we give a physical account of this connection that does not simultaneously apply to irrelevant phenomena? As far as we can tell, every electron "matches" every other electron in the universe, but we do not say that every electron is "about" every other electron.

Suppose a scientist monitors my brain activity while I think about what an atom is. The scientist may identify the parts of my brain that become active during the process and must be functioning properly in order for the thought to occur. But no matter how exhaustive this monitoring, it will reveal biochemical events that have no greater resemblance to an atom than do many events or states that are clearly not about atoms. Any single atom of my brain (or of my left foot, for that matter) is a better physical "match" for the atoms outside my body than is the biochemical sequence that occurs when I think about what an atom is. And we cannot trace a physical, causal connection between the sequence and the object of the thought, which in this case is not a specific atom existing at particular time and place. Because of difficulties such as this, in cognitive science the biochemical sequence is referred to as the "correlate" of the mental event that accompanies it. The term "correlate" is used because the physical sequence and the thought have different properties. Correlation is not identity. The fact that the pressure of a gas correlates with its temperature at constant volume does not mean that pressure is reducible to temperature.

When I think about what an atom is, the thought's intentionality directly connects it to every atom that has ever existed or ever will exist. But no scientist will be able to find direct physical connections between the correlate of the thought and every atom that will ever exist; he will not even bother to search for them. So the cognitive scientist is left to speculate that the mental conception somehow "supervenes" on the correlate.

In desperation, we might try to identify a proposition or thought (or intentionality itself) with its physical correlate by using the concept of representation. We could say that the correlate represents the thought in roughly the way that a sentence does. But to represent an object is not to be that object. The word "green" can usefully represent the color green, but it is not in fact the color green and does not have the attributes of a color. The written word "green," being composed of letters, must conform to a certain pattern of geometric shapes, while the actual color of green has no specific shape. Moreover, I can write the word "green" in black ink on white paper, as it need not have any visual "greenness" whatsoever. We cannot reasonably argue that because the word "green" need not appear to be green, the same can be or must be true of the actual color.

A primitive computing device, the abacus, provides another example of representation. Our mind is capable of seeing the positions of beads on the wires of the abacus as representing numbers. But the properties of numbers, it hardly needs saying, are different than the properties of beads on wires. For instance, beads on wires cannot said to be integers. Confusing a representation and its referent is a fallacy common in magic and superstition, exemplified, for instance, by the voodoo doll.

The distinction between representation and identity undercuts Carrier's related claim that "since one does not need anything more than physics to have communication and computation, you do

not need anything more to have logical laws." The fact that physical objects and events can represent thoughts for purposes of communication and computation does nothing to demonstrate that thoughts are physical. Even so, Carrier leaves no doubt that this is exactly what he is arguing. He later claims:

> Both inductive and deductive logic and all mathematics, including probability theory, can all be represented and carried out by an appropriately-organized assembly of [logic] gates, and therefore all such procedures are mechanically reducible to them (or their functional equivalents--i.e. the functional equivalents of the gates themselves, or their various assemblies).

It is easy to verify further that representation of logic with "logic gates" (i.e., staged electronic switches) does not mean that logic is reducible to them. We can usefully represent the relative planetary movements of our solar system with a model that consists of plastic balls moved by belts and pulleys driven by an electric motor. Such a model can function as an analog computer that predicts the future positions of the planets. But the fact that we can model planetary movements electromechanically does not demonstrate that the solar system itself is an electromechanical system as opposed to an inertial-gravitational one, or that anything that we would recognize as a true planetary system could be electromechanically driven.

The nature of an idea or proposition is such that we cannot conceive of it as a physical object or state. My idea of green apple is conceptually transparent; it acts as a kind of lens through which I can interact with objects and form mental and verbal propositions as to whether or not they are green apples. Through it I can connect with objects that no longer exist or that do not yet exist, namely, all those that were or will be green apples. An idea could not function that way if it were tightly and opaquely confined in time and space in the way that a physical event is. Substituting the term "virtual model" for "idea" does not change matters. Are the properties of a "virtual model" the same properties that we could observe or detect in a sequence of synaptic discharges in the brain? If not, then we cannot identify a "virtual model" as such a sequence.

Mental Causation

As we have seen, the intentionality of our thoughts does not fit easily into a naturalistic philosophy that seeks to give explanations exclusively in physical terms. For a starker inconsistency between naturalism and reason we must go on to the related topic of mental causation.

Here is a synopsis of this part of the argument and Carrier's response. C. S. Lewis said that if we arrived at certain conclusions purely for physical, biochemical reasons then we would be excluded from arriving at those conclusions for logical reasons. Since we cannot repudiate our own rationality or power of reason, we cannot attribute that power solely to physical processes in the brain. While Victor Reppert in *CSLDI* identifies six different subarguments or aspects of the AfR, this one is its core. The standard naturalistic reply is that chemical reasons do not exclude logical reasons. But Reppert insists that given a realistic view of science and nature, some reasons or explanations do indeed exclude others, and the contradiction that Lewis identified is genuine. Carrier replies that what happens inside a computer occurs both for physical reasons and for logical reasons (since computers can give logical outputs) and therefore the same can be true in our brains, particularly since brains and computers have many similarities.

The computer-brain analogy, if not the strongest reply to the AfR, is surely the most vivid. And if we do not press the comparison unduly it illustrates naturalistic functionalism. Put simply, that philosophy says that what functions as a mind is a mind, with the assumption that the functions involved can take the form of physical changes in physical systems. To understand

why it cannot rescue naturalism, I will frame the AfR in terms of four corollaries--two of naturalism and two of reason itself. The corollaries of naturalism are as follows:

First Corollary of Naturalism

> To the extent that changes in natural systems have causes, those causes are potentially available to the senses either directly or by means of scientific instruments.

Second Corollary of Naturalism

> Every belief accompanies a natural state, and the properties of a belief are wholly dependent upon and determined by the natural state that it accompanies.

Although these corollaries are my own, I believe that they faithfully capture relevant concepts and will prove difficult for proponents of naturalism to disown. The corollaries imply that all causes of belief are potentially available to scientific observation, measurement, and description. They also assume causal closure. Under closure, to the extent that a natural state is caused, it is caused solely in accordance with natural law. We will be comparing these with two corollaries of reason as follows:

First Corollary of Reason

> Reason includes, although it is not limited to, the acceptance of a belief due to the accurate, conscious perception that true premises logically entail it.

Second Corollary of Reason

> A belief may be considered to be held rationally only to the extent that what are consciously perceived by the holder to be the reasons for his accepting the belief are in fact the reasons for his doing so.

Our first corollary of reason identifies one species of reason as the process of arriving at a belief due to the conscious perception of logical grounds. "Perception" here describes an introspective act[3]. The logical entailment identified in the first corollary of reason is not available to the senses, which disqualifies it as a cause of change in a natural system under the first corollary of naturalism.

At the website naturalism.org several tenets of naturalism are listed. Under the first heading, ***What exists***, we find the statement, "Nothing about us escapes being included in the physical universe, or escapes being shaped by the various processes--physical, biological, psychological, and social--that science describes. On a scientific understanding of ourselves, there's no evidence for immaterial souls, spirits, *mental essences*, or disembodied selves which which stand apart from the physical world [emphasis mine]." The problem with this claim should be evident. Measurement, which is the basis of scientific description, cannot be carried out on that which is only accessed introspectively. Realities that are not available to the senses obviously "stand apart" fundamentally from those that are.

Under the second heading, ***What constitutes knowledge***, we find this: "Scientific empiricism has the necessary consequence of unifying our knowledge of the world, of placing all objects of understanding within an overarching causal context. Under naturalism, there is a single, natural world in which phenomena arise." Whatever can be perceived only through introspection

cannot be fitted, however, into the same causal context with that which is accessible through the senses. For example, we may find that one thought causes another by association or suggestion. But it is meaningless to say that one physical event causes another by association. Suppose that a spark causes a fire. The spark may be described as a fragment of rock or metal at high temperature and the flame as the combustion of specific materials. The physics of the event can be described in finer and finer detail. But the elaborations have the same quality as the original description. We can see how the fibers comprise the threads of the explanation, so to speak, and how the threads in turn make up the cloth. But successions of chemical states and successions of thoughts are too dissimilar in their apparent properties to be held within a single causal context, and this dissimilarity is the reason why they are perceived by different means.

We can make the claim (tentatively, at least) that the properties of thoughts that allow thoughts to lead to one another through association do not actually have the causal power that such properties seem to have. This allows us to avoid the impossible task of holding two utterly different types of causation within a single context. We can then propose that brain chemistry alone is the causal engine, while causation by association is a benign illusion. But whether or not this tactic works in the case of association, we cannot resort to it where reason is concerned. If it is illusion that one thought can cause another by logical implication, then most of our knowledge--including virtually all of our scientific knowledge--is, in fact, mere belief.

I believe that Carrier's argument regarding mental causation, and naturalism itself, falters precisely on this point. Carrier says about perception:

> So only the second prong is left: the causal foundations of perception. But to challenge that is to challenge consciousness itself. For there is no significant difference between a veridical visual experience and a veridical rational experience. If any machine can create the former, it can create the latter, just as it can create one for sound, touch, emotion, taste, smell, pain, pleasure, panic, and so on, whatever chance and advantage should produce a perceptual system for (the human sensory array is by no means exhaustive: bats see sound, lobsters feel magnetic fields, and so on). So Reppert cannot single out reason here. An attack on rational perception is automatically an attack on all perception of whatever kind.

The "significant difference"--the crucial difference, in fact--between visual experience and rational experience is that the former is sensory and the latter is not. When I say that I "see" a logical conclusion I am not referring the sensation of an image on my retina but to an act of understanding. It is not an impingement of a physical state upon my ears, eyes, nose, or tongue, nor can it be categorized with any of the multitude of bodily sensations my brain can process. Least of all does such a perception amount to one part of my brain feeling or "seeing" another part of it. Remember, no part of my brain can be identified with what it is that I rationally perceive. What one part of my brain may sense in another part is a representation rather than the representation's referent. Suppose the referent is a proposition (e.g., "All men are mortal"). All that my brain can be imagined to sense is the representation of the proposition (e.g., words scribbled on a notepad), but what I rationally perceive is the proposition itself (that "all men are mortal").

For the reasons I have referred to, it would be a fallacy to think that we make sensory contact with a proposition, such as an argument, because we can read it or hear someone explain it. Again, the written or spoken form of a sentence is a representation. We make sensory contact with a representation, which prompts us to perceive an argument, proposition, or idea. In that loose sense we hear or read the argument. But strictly speaking, we do not make sensory contact with an argument in the way that we do with a chair, the clang of a bell, or even--by means of a Geiger counter--the decay of an atomic nucleus.

Because representations are physical, they can be made to do work in a physical system. For example, in the early days of computing data was represented by holes in punch cards. The holes engaged small cogs or contacts, which then made other things happen in the computer through physical cause-and-effect. At the end of this process new representations were generated, such as printed addresses on fanfold paper. Computing has become more sophisticated, with electronic switches in microcircuits replacing holes and cogs, but the nature of the process has not changed. Physical representations of data are input, causing physical changes to occur in sequence. At the end of this causal chain other representations are output.

In the case of the human mind or brain, physical representations of propositions are input just as with a computer. After that, however, what do the work of causing a belief are (sometimes, at least) *the propositions as such*. If physical representations of propositions--as opposed to the propositions themselves--do all of the cause-and-effect work of generating my beliefs, then necessarily my beliefs are not generated by propositions. The rationality of all of my beliefs would be compromised under the second corollary of reason, above. The belief that only natural causes can affect the state of the brain, for example, could not be caused by arguments about the sufficiency of natural causes.

To confirm the intractability of this difficulty for naturalism, recall from our previous consideration of intentionality that the aboutness of a sentence cannot be identified with the sentence's physical properties. Yet those properties are precisely what must do the work of generating states in a natural system. We can summarize:

1. Only the physical properties of representations can generate functional states in computational systems.
2. Propositions cannot be identified with the physical properties of their representations.
3. Therefore, propositions cannot generate functional states in computational systems.
4. Propositions generate some beliefs in minds.
5. Therefore, all beliefs in minds cannot be identified with, or wholly depend upon, functional states in computational systems.

Perhaps there is an escape hatch for naturalism regarding statement No. 1. If the representations it refers to are caused by propositions then it can still be said that propositions can cause functional states, albeit indirectly. However, this gambit requires us to back up in the causal chain to some system that is not merely computational, that is, to a system in which propositions rather than just representations do causal work. We are faced with infinite regress if we merely back up from one computational system to another. So far the argument has no obvious naturalistic solution[4].

Here we will pause to take a closer look at the second corollary of reason set forth above. The AfR as I have outlined it is strengthened by that corollary even though it will stand without it. The second corollary of reason states that rationality depends upon correctly perceiving the grounds for beliefs.

One might argue that we can rationally believe a proposition without perceiving good reasons for our beliefs. All of us believe some scientific propositions, for instance, that we are incapable of justifying in detail. To hold a belief on the basis of authority does not of itself disqualify a belief as rational. Most of us can offer good reasons for generally trusting, not just the consensus pronouncements of science, but also other authoritative opinions for which we cannot offer detailed explanations.

What about beliefs based on guesses and intuitions? These may be truthful even when cloaked in rationalizations that the holder mistakes for logical grounds. Suppose a juror perceives the evidence of a criminal defendant's guilt, as presented by a prosecutor, to be more conclusive

than it actually is because of having subconsciously picked up clues from the defendant's body language and speech patterns. Can't the juror's belief be considered rational, arising as it does from the guilt of the defendant? I do not believe that it can unless the intuition is made the subject of rational reflection or discussion. The juror may say to himself or others, "Something about the defendant makes me feel that he is untrustworthy, and I have learned from experience that my feelings about people are usually right." In such a statement we would recognize a higher degree of rationality than we would in the unreflective--and incorrect--insistence of the same juror that the formal evidence has strength that in reality it lacks.

Or imagine that we try to persuade someone else of the truth of a belief that we hold. We present what we think are good reasons but the person remains unconvinced. Perhaps the person replies that our case is weak, that our arguments are inconsistent, or that reasons weighing against our position seem to be stronger than those in its favor. These responses allege that we are mistaken, but they do not impugn altogether our rationality in holding the belief. But suppose instead that this person claims that we do not hold the belief for the reasons that we think we do. We would be foolish to insist that even if our opponent were right, we might nevertheless be rational in maintaining the belief. Knowing the reasons for our beliefs is part of what it means to have reasons, and having reasons for believing is part of what it means to believe rationally.

It is conceivable that few of our beliefs are purely rational, even though many of them are true. Perhaps some that have little rationality are nevertheless in some mysterious way worthy of belief. But it is difficult to deny that what we rationally believe is rational only insofar as we discern why we believe as we do, and this fact reinforces the AfR.

What I have outlined to this point is an updated version of Lewis's argument cast in computational terms. It refutes the second of our two corollaries of naturalism, but only to the extent that belief states cannot be wholly dependent upon natural brain states. The door is still open for natural states to be necessary conditions for rational beliefs, at least in creatures such as human beings. The findings of cognitive and neuroscientific research continue to have explanatory value, but it is conceptually impossible for science to explain human reason completely in terms of biochemical states. We will return to the relevance of scientific explanations again below.

In *CSLDI*, Victor Reppert makes a similar argument from mental causation by referring to the propositional content of mental states. He says that if naturalism were true then the propositional content of one mental state could not be the cause of another state such as a rational belief. Since rational beliefs must be caused by the propositional contents of earlier mental states, naturalism cannot be true. Carrier's responses to the argument are revealing. First, he attempts to associate propositions with virtual models, which as we saw previously he considers to be physical:

> Propositions are formulated in a language as an aid to computation, but when they are not formulated, they merely define the content of a nonlinguistic computation of a virtual model. In either case, a brain computes degrees of confidence in any given proposition, by running its corresponding virtual model and comparing it and its output with observational data, or the output of other computations. Thus, when I say I "accept" Proposition A this means that my brain computes a high level of confidence that Virtual Model A corresponds to a system in the real world (or another system in our own or another's brain, as the case may be); while if I "reject" A, then I have a high level of confidence that A does not so correspond; but if I "suspend judgment," then I have a low level of confidence either way.

We have touched on the problem with such a claim. Consider a physical object, say, a ship. A ship can be made in various shapes and sizes out of various materials, but it has certain defining characteristics that can be seen and measured, and a function that can also be measured. So it is with a tree, a thunderstorm, or a chemical reaction. All natural objects, events, or states have features, sometimes even emergent features, which can be sensed or detected and measured. And the features that we can detect and measure make the object what it is in natural terms. So, what detectable features make a virtual model what it is, and what is the measurable content of such a model? If that which makes a virtual model what it is can only be perceived by introspection, then under naturalistic assumptions it cannot cause changes in a natural system.

Later in his essay, Carrier specifically cites and then addresses Reppert's argument about the propositional content of mental states as follows:

> 6. The state of accepting the truth of a proposition plays a crucial causal role in the production of other beliefs, and the propositional content of mental states is relevant to the playing of this causal role. [Quoting *CSLDI*.]
>
> Brains are computers. As such, the output of one computation (including the output of confidence level) is often physically the input of another computation, and it thereby has a causal effect on that other computation's output. Every conscious computation in the brain is the computation of either a virtual model or data physically connected to or computed from a virtual model (such as a confidence level). Since a proposition literally is the content or output of a virtual model, propositional content therefore literally has a physical-causal effect on further computation that relies on that virtual-model computation (which literally is the "proposition" in question).

The claim here that "a proposition literally is the content or output of a virtual model" is remarkable. Elsewhere in his essay Carrier makes it clear that virtual models are not just constructed by the brain but by actual computers as well[5]. And he also leaves no doubt that by "literally" he also means "physically," since he says that output is "often physically the input of another computation." So a proposition is (or at any rate can be) physically and literally the output of a computation. Let's say that the output of a computation is the printout of a proposition. Physically the output is carbon (toner) and cellulose (paper). So the natural properties of carbon and cellulose must be, in some respects, the natural properties of the proposition. Worse yet, if the proposition is the same as the meaning of its representation, and if the proposition consists of carbon and cellulose, then the meaning of the representation must consist of carbon and cellulose. This makes no sense. We can substitute electrical impulses in circuitry for carbon and cellulose and it still makes no sense.

We will be generous here and take Carrier to be saying, not that a proposition literally is the physical output, but that it is realized in the physical output. A triangle, for example, can be realized in wood or plastic or glass. A ship, as stated earlier, can be realized in wood or steel. Perhaps a proposition can be realized in carbon and cellulose. On second thought, that won't work either. A triangle cannot be realized except as an object with triangularity. The English word "triangle," by contrast, has no triangularity. Likewise, to be the realization of what a ship is, an object must be able to float or be intended to float. It must be sufficiently shiplike. But the word "ship" need not be shiplike at all to represent what it does. Representation, therefore, does not entail realization, much less identity.

Carrier's attempt to naturalize propositions or propositional content must be judged a failure, and as a consequence the bulk of his extensive analysis collapses. It cannot be salvaged by his many references to scientific opinions about human cognition, nor by his frequent criticism of Reppert for failing to pay sufficient attention to such information. The AfR is not, as Carrier

implies, an argument from scientific ignorance. It should be apparent by now that the AfR identifies logical constraints on our understanding of the relevant scientific facts.

A paper written in 1977 by two cognitive researchers, Richard Nisbett and Timothy Wilson, reveals the conceptual limits on scientific analysis of cognition[6]. Nisbett and Wilson surveyed experiments in which test subjects were monitored for their reactions to various controlled situations and then afterward asked why they had reacted in the way that they did. Evidence showed that in a wide range of circumstances people had reacted for reasons different than those which they later reported. Some of these situations called for problem solving and reasoned judgment. For example, experiments have shown that the willingness of someone in a room to come to the aid of someone else in an adjoining room who sounds as if they are in distress diminishes as the number of bystanders increases. Subjects in a crowded room are typically slow to offer aid, but when questioned later they deny that the number of people around them had anything to do with their decision not to act. Other studies found that ratings of stocks by brokers, and diagnoses of illness by therapists, surprisingly poorly reflected the relative importance of the various criteria according to which they reportedly made their judgments. The article contains a list of other similar findings.

Nisbett and Wilson go so far as to state, "The evidence reviewed is then consistent with the most pessimistic view concerning people's ability to report accurately about their cognitive processes"[7]. Elsewhere in the article they moderate this claim, but think about what it would mean for us to accept it without qualification. Or more properly, think of how impossible it is for us to do so. What happens when we apply this assessment recursively to the judgments of Nisbett and Wilson themselves? If we were to ask the authors why they arrived at the conclusions expressed in their paper, presumably they would say that they were led to them by rational analysis of the experimental results they cited. But if we had grounds for doubting their self-report then the reasonableness of the paper's conclusions would be open to question. We could ourselves check to see whether the evidence in the paper supports the judgments it contains, except that absolute skepticism about cognitive introspection would cast doubt on our own judgment as well.

Nisbett and Wilson could argue that rigorous scientific methods, measurements, and records can correct for our failures to appraise accurately our own mental processes. However, the belief that scientific tools can play a corrective role is itself the kind of judgment that will not survive a universal indictment of introspective knowledge. So there is a limit to how much skepticism science can ever engender regarding our perceptions of our own cognitive processes. We cannot judge as reasonable a finding that denies us the ability to make reasoned judgments, or even one that under logical analysis is found to entail such a denial.

Epiphenomenalism and Volition

Although this paper was written specifically in response to Richard Carrier's review, there are points that Carrier does not touch on that are important enough to the AfR in general to warrant attention here. In support of C.S. Lewis and Victor Reppert, I have argued that while rational processes are conditioned by physical activity in the brain, they cannot be fully explained by it. This implies that when the right physical conditions are present, something nonphysical can occur. We can illustrate this with a compass. When the pieces of a compass are assembled carefully, the needle begins to be acted upon by a force that cannot be detected during the assembly process. Admittedly, tough questions need to be answered by anyone who proposes that nonphysical, rational processes have physical effects on natural systems such as brains. But first we must consider a philosophical position, epiphenomenalism, which allows for genuinely rational beliefs but denies that they have any physical effects. Epiphenomenalism is too pessimistic a philosophy to be popular, but refuting it will give us more reason to believe that interaction between the mental and physical sides of reality does occur.

Consider a person's act of blinking when an object comes flying toward his or her face. Because eyes are delicate enough to be injured even by light impacts, closing the eyes quickly when a projectile comes toward them is a rational thing to do. I may acknowledge the rationality of blinking under such circumstances and even form the intention to do so. When a spray of liquid from a pressurized soda can comes suddenly at my face, however, I will blink reflexively. The correspondence of my act of blinking to my belief that it is a reasonable thing to do and even the harmony between my blinking and my intention to blink do not change the fact that the blink occurs for reasons other than my belief and intention. Epiphenomenalists propose that what is true of my blinking is true of all of my behaviors without exception. It may seem as though my beliefs and intentions cause my actions, but that is a misperception. While I am consciously believing and intending, the unconscious but sophisticated computational processes of my brain are making my arms and legs move, my mouth speak, and my fingers punch the keyboard.

Epiphenomenalists admit to two great, probably insoluble, mysteries. The first is conscious experience itself. The second is the alignment of conscious experience with behavior. But having accepted these two mysteries, they avoid others. They need not assume the Herculean labor of naturalizing beliefs, desires, purposes, and intentions. Neither must they struggle to maintain closure of physical causes against problematic interference from the mental side. Technically, epiphenomenalism is dualistic because it separates nonphysical experience from physical behavior, but it is a tame dualism that does minimal damage to naturalistic assumptions. The logic of epiphenomenalism, though hard to accept, is easy to grasp. Any behavior can conceivably be programmed, but it is not clear how beliefs, desires, and intentions can be programmed[8]. I can program a computer to respond to an e-mail, but how can I or anyone else program the computer to respond intentionally rather than unintentionally? If the survival of the computer depended on a response that programming was sufficient to ensure, what role is left for intentions to play? Substitute organism for computer and the argument is complete.

The fatal flaw in epiphenomenalism lies in the experience that we all have that we intend some actions and not others. Deciding to read a book as opposed to watching television is intended or volitional, while blinking at a sudden spray of soda is not. Epiphenomenalists' only explanation for the experience of volition--that it is an illusion--must be examined critically. Coherence is the minimum requirement for any philosophical conclusion, and to determine coherence in the present instance requires a brief analysis of the relationship between experience, conceivability, and reality.

In his *Treatise of Human Nature*, the philosopher David Hume observed that experiences take the form of impressions that can be simple or complex[9]. His example of a simple impression is the sight of a patch of red color. A more complex impression, he says, might be the collage formed by the sights and sounds of a city. He notes that complex impressions might be of objects that do not exist. He cites the example of the New Jerusalem described in the Bible, a city with golden streets and buildings made of jewels. How is it, Hume asks, that he can form an impression of something that he has never seen? Hume goes on to show that the answer lies in the complex nature of an imaginary construct. There are such things as city streets and such a thing as gold, and by combining them we may form a mental picture of something that does not literally exist. Another example is the unicorn. There are no horses with horns growing out of their heads, but because there are real horses and real horns we may combine our impressions of them into a mythical beast.

What is true of imaginary constructs is true of illusions. An example is the "moon illusion," in which we misperceive that the image of the moon occupies a larger area of the nighttime sky when it is near the horizon than when it is overhead. Differences in size both of visual images and of objects as such can be real, and objects in the sky can indeed be seen, but the moon illusion incorrectly conjoins these impressions.

An intriguing fact about impressions--what today might be called percepts--is that as we move from complex instances to simple ones it becomes progressively harder to conceive of the possibility that they are unreal. Look at the following statements:

1) Brown horses with natural white spots on both ears do not exist.
2) Brown horses do not exist.
3) The color brown does not exist.

In the first case, if we seem to remember a brown horse with white spots on the ears we might be mistaken. The second example is harder to contemplate, but perhaps a rare, undiagnosed ocular dysfunction has misled us into seeing black and gray horses as brown. With the last statement we have reached the frontier of coherence. It is inconceivable that the color brown does not exist, but why? Had I been born profoundly color blind I would know brown only at second hand. The claim that brown does not exist, while less than plausible, would be conceivable for me. Or say that brown, instead of being a common characteristic of the visual landscape, were a fabled color so pleasing to the eye that it becomes the favorite of anyone who looks at it. In that case our knowledge of brown would owe to a complex idea combining the simple impressions of sight, color, pleasure, and relation. "Brown" would then be like "unicorn" and we could entertain not just the possibility but even the likelihood of its nonexistence.

As it is, most of us are acquainted with brown by sight. Our idea of it owes to a simple impression, as Hume would say, which is why a denial of its existence defies conceivability and falls short of coherence. We can draw an analogy to language, specifically to words and propositional statements. A statement can be judged as false because it can be broken down into simpler units that convey meaning faithfully. Once we reach the smallest semantic units, the individual words, we can cast doubt on them only by impugning our ability to assess the truth or falsehood of the statement they comprise. Take the statement, "One of the words of this sentence does not have the meaning that it appears to have." This claim leaves us in an impossible position. If the statement were true, it could not mean what it appears to say, which would then mean that it is not true. Internal dissonance of this kind is called self-referential incoherence. Famous examples are the liar's paradox ("I never tell the truth") and Russell's set paradox[10].

Because we know the color brown from a simple impression, the claim that brown does not exist is incoherent. It asks us to draw upon our knowledge of brown while denying the basis of that knowledge. Note, however, that although the claim that brown is altogether illusory cannot be true, the possibility remains that the apparently brown color of an object may be illusory in a particular instance.

The reason for this analysis of the complex nature of fantasies and illusions is that epiphenomenalism, as we saw above, denies that mental states as such can have physical effects. Nevertheless, our knowledge of the difference between intended and unintended actions comes from a simple impression; therefore its denial is incoherent. Whether we label the object of this impression as "will" or "volition" or "agency," it cannot be broken down into perceptual pieces that our minds could have incorrectly stitched together. But just as with the existence of the color brown and brown objects, our present argument leaves open the possibility that the deliberate nature of our actions may be illusory in particular instances.

Since volition must be real, the epiphenomenalist's argument leads in an unexpected direction. Programming might be sufficient to explain behavior divorced from experience but not behavior as related to experience. This can only mean that our experience is incompatible with a purely computational or even functional model of mental processes. This is the same conclusion that the AfR yielded in the preceding section. We now have two independent lines of argument both showing the inadequacy of a naturalistic account of mental activity.

The AfR, Causal Closure, and Physics

Granted that mental activity is dependent upon but not identical to chemical reactions in the brain, how can it be imagined to influence those reactions? A majority of philosophers doubts that mental processes can, without being physical, exert physical influence. To understand why the majority is wrong on this issue, return for a moment to the work of David Hume. Part of Hume's continuing legacy is his demonstration that physical cause and effect do not carry the force of logical necessity. If I light a match and touch it to my finger, it is a matter of logical necessity that the flame cannot simultaneously burn my finger and not burn my finger. But that the flame does burn my finger is not logically necessary. The burning of my finger is an event that happens to occur but does not have to occur, at least not out of logical necessity.

Hume's discovery, which we will discuss further below under the problem of induction, has survived every attempt at refutation. The philosophy of science as a whole continues to accept that physical causes are not linked to their effects by logical necessity. Hume made this discovery purely though reflection, however. No one, least of all Hume himself, anticipated that science would one day provide support for his conclusion. It was able to do so because Hume's analysis of cause and effect implied the possibility of physical events occurring in the absence of physical causes.

In the realm of philosophy, Immanuel Kant failed to embrace Hume's radical views on causation and thought that the mechanistic physics model of Newton was logically inevitable. If relativity undermined this assumption, quantum theory demolished it. Quantum physics stunned the scientific community of the early twentieth century. Karl Popper, the prominent philosopher of science, struggled with it. Einstein famously rejected it, even though he helped to lay its foundations. The significance of quantum physics to our present discussion is not that it provides a mechanism by which mental events can affect natural processes, but that it undercuts all arguments that implicitly give physical causation the status of logical necessity.

Many readers will be familiar with the basics of quantum theory. I invite those who are not to imagine a large barrel full of liquid with a small hole near the bottom. If we know the size of the barrel, the size of the small hole, the density and viscosity of the liquid, and other physical variables, we can calculate with a high degree of precision how much liquid will leak out of the hole in a given period of time. What is true of the barrel is roughly true of a lump of uranium 238, if we substitute the composition of the atomic nuclei for the characteristics of the fluid and refer to a flux of radiation rather than to leakage. Both of these examples confirm our common sense view of cause and effect because they exhibit uniformity. The leakage rate will be similar for two similar barrels of fluid, and likewise the flux of radiation from two comparable lumps of uranium.

The strangeness begins when we scale down to two atoms of uranium 238 instead of two lumps. In spite of the atoms having identical compositions, one may decay in the next minute while the other remains undecayed for the next million years or more. Common sense tells us that there is some undetected difference in the two atoms that accounts for their different times of decay, but common sense is wrong. How can this be?

To find out, let's turn from barrels to photographs. We take a photograph and in looking at the print see that edges of the image are not entirely crisp. Was the focus incorrect? Did the camera move? We photograph the same object over and over and make estimates of the possible camera movement and focal errors, but neither these nor imperfections in the film account for all of the fuzziness. On the other hand, we can still easily determine the shape of the object in the photograph, which is relatively large. As we photograph smaller and smaller objects, the fuzziness occupies a larger portion of the image until nothing is visible but a blurry smear. Does the extremely small object have a distinct shape that does not register, or does the

information consisting of the object's shape simply not exist? This illustrates, in an admittedly rough way, the problem that early twentieth-century scientists encountered as they tried to gather information about subatomic particles.

To continue with our photographic illustration, suppose we determined that the fuzziness of our image was not the result of technical deficiencies. Suppose instead that it proved to be theoretically unavoidable when registering an image by reflection of energy. This discovery would be doubly disturbing because the concept of shape is impossible to divorce from registration. If an object can never in principle register a shaped image then we are forced to conclude that the object is shapeless. Likewise, the vast majority of physicists have concluded that information we routinely find in large objects and states is lacking at the subatomic level. Mysteriously, information emerges as a statistical effect; shapeless smears when assembled acquire shape.

In the case of radioactive decay we have a smeared time image rather than a spatial one. We can know the exact time of decay of a nucleus by probing it with enough energy to cause its decay, but that would amount to forfeiting information in the process of gathering it. If we wait for the nucleus to decay spontaneously, information pinpointing the time of decay will come into being only when decay occurs. The undisturbed composition of a single nucleus contains only a fuzzy or probabilistic indication of its decay time. However, when trillions of atoms of uranium are assembled, the time image of decay assumes a definite shape.

Quantum theory challenges our imaginations. The pioneering physicist Neils Bohr famously quipped that whoever is not shocked by it has not understood it. Nevertheless, the vast majority of physicists subscribe to it in all its strangeness. They do so not because they are inclined to believe the absurd but because an ever-expanding body of experimental evidence, backed up by rigorous mathematical modeling, has compelled their assent. One of its implications is that an information deficit exists at the root of physical reality. Can this inadequacy of physical causation be supplemented by mental causation? Yes, but probably not in a way that directly supports the AfR. If the brain were a kind of Geiger counter or amplifier of quantum events--which it likely is not--randomness might result, but not rationality.

The contribution of quantum theory and its partner field, chaos theory, to the AfR is to demonstrate that all physical systems cannot for the sake of argument be put in a mechanistic straightjacket. As a further example, consider quantum tunneling, a process that allows the law of conservation of energy to be compromised slightly to allow shifted outcomes. Imagine a soccer player and a ball at the foot of a small hill. If the soccer player kicks the ball with just enough force, it will store potential energy as it rolls up the near side of the hill and then "spend" the stored energy rolling down the other side and several meters out across flat ground. If the player imparts a bit less energy to the ball it will store energy while it rolls up the hill, then stop, then spend the stored energy rolling back down the near side. But what if the weakly kicked ball, as it begins to slow on the near side of the hill, could "borrow" some of its stored energy in order to reach the crest? After reaching the crest, the weakly kicked ball could pay back the borrowed energy by rolling down the far side of the hill with slightly less speed and traveling a shorter distance out across flat ground. Given the possibility of such borrowing, with exactly the same amount of energy the ball could end up either on the near or far side of the hill--two different outcomes.

In the case of quantum tunneling, the borrowing occurs just under the theoretical threshold of observation, but experiments yield results that can be explained in no other way. In the case of the brain, the complexity of neural events may create a chaotic threshold high enough to prevent observation of a kind of energy borrowing that allows mental causation to affect physical outcomes. Naturalists argue that some physical explanation or other of biological processes, including those in the brain, have thus far proved adequate, and that no violations of physical laws have been observed in these processes. The arguments from reason and volition,

however, describe an absolute theoretical limit to the adequacy of biological explanations. Further, reason and volition only begin to come into play as attendant physical systems become staggeringly complex.

The simpler the system being studied, the easier it is to verify that no physical anomalies are occurring as it operates. At the same time, simplicity makes it harder to conceive of reason and volition being present. It is hard to imagine that a lobster does anything intentionally (rather than unintentionally) or that it ponders logical relations. It becomes progressively less difficult to imagine some role for these phenomena when moving from a cat to a dog to a chimp to a four-year-old child. By the time we get to an adult human of normal intelligence, we are faced not just with the most complex organ known to science--the human brain--but the most complex integrated system of any known kind. Perhaps not coincidentally, the increasing complexity of the neural machinery as we move up the scale makes it more difficult to resolve fine-grained processes in such detail as to rule out small and fleeting anomalies that may be occurring.

Granted that there is an argument for mental affects, what would a physical event caused by a mental event look like? It would look like an event for which sufficient physical cause is lacking. The mental cause would be invisible to sensory observation. The only way to "see" the mental event would be to experience it from the perspective of the mind in which it occurs. We do, as it happens, experience causally effective mental events. And physical effects for which physical causes are inadequate, though strange to contemplate, are not beyond reason.

The Laws of Logic

One of the subarguments of the AfR that Victor Reppert includes in his book has to do with the difference between the laws of physics (or, laws of nature) and the rules for rational thinking, referred to as the laws of logic or laws of thought. If mental processes form a subset of physical processes, why do different rules apply to each? Richard Carrier replies that the laws of logic are really physical laws:

> So also for the law of non-contradiction--that "law" is just like any other physical law: it describes any universe that contains distinctions (and causal and ontological consistency), which is (I suspect) every universe except the null set, since we cannot conceive of (compute) any other universe without positing at least one consistent distinction. This, again, is not because of some mysterious logical superlaw, but because the physical facts of any universe we care to compute are just so.
>
> In short, the laws of logic literally are the (potential or actual) physical properties of every (potential or actual) universe that they describe.

Carrier's opinion notwithstanding, I invite readers to verify that books on physics (and science generally) do not list the laws of logic among physical laws. The reason is that the laws of logic include truth as an objective value. There is a certain way thought must proceed if it is to reach truth. Scientists do not ask about the truth-sensitivity of natural processes. Neither the electrons in the outer shell of an atom nor the outermost planets of the solar system are farther from the truth by virtue of their placement.

Above I pointed out that the meanings of propositions cannot be identified with the physical properties of sentences. It doesn't matter whether these are the properties of a sentence on a certain piece of paper (token properties) or the properties of the pattern of words and letters apart from any particular occurrence (type properties). Since the truth of a proposition relates only to its meaning, its truthfulness cannot be identified as a physical property. Therefore, rules

that depend upon truth as a value cannot be identified as a subset or special application of rules that pertain to physical properties.

Carrier is correct to the extent that in any physical universe where mental activity occurs, the laws of logic will apply. That observation by itself does not refute Reppert's argument, however.

The Problem of Induction

Carrier's suggestion that physical reality is somehow rational, however inappropriate it is within his argument, does touch on a real truth. It happens to be the same truth that he carefully avoids in his remarks on the problem of induction:

> Of course, even following Hume and Reppert, inductive reasoning is fundamentally circular for everyone, even for God. Everyone is in the same boat, theist or atheist. All one can do is posit a hypothesis (e.g. "God exists" or "Nature is consistently uniform") to explain the evidence (consistent uniformity) and then constantly test that hypothesis (so far, it has never been falsified, and has been abundantly verified). To accept this line of reasoning, all one need do is reject the groundless and self-defeating methodological principle "Maybe, therefore probably." And the theist is in no better position than an atheist here.

Put concisely, the problem of induction states that all of us, scientists included, assume that the universe is intelligible. Without this assumption, neither scientific conclusions nor common-sense expectations about the world could have any value. Carrier believes that statistical reasoning holds the key to solving this problem. It is true that subtle links exist between deductive and inductive reasoning. Nevertheless, the fallacy of statistical approaches to the problem of induction was recently pointed out by John Foster[11]. As Carrier himself admits, a consensus of professional philosophers does not believe that anyone has solved the problem[12].

Against Carrier, nature's intelligibility *does* put the theist in a "better position." The two propositions, (1) that matter is derived from self-existent mind and (2) that mind is derived from self-existent matter, are not symmetrical or equally coherent. Mind, or "thought" to be precise, has a native intelligibility that matter lacks. Wherever matter behaves intelligibly it must, as far as we can tell, draw its intelligibility from thought. Matter's lack of inherent intelligibility--the fact that physical cause-and-effect lacks the force of logical necessity--is actually the fundamental issue underlying the problem of induction.

An easy way to understand the problem of induction is to contrast patterns with rules. Most readers understand that a pattern can owe to mere chance and that when it does it has no predictive value. For example, if in tossing a coin we happen to get a string of three or four heads in a row, we have generated a short pattern of repeating heads. But the pattern would not increase the odds of getting heads on the next toss. Likewise, if we happened to get the same number several times by rolling a die, the odds for the next roll would remain unchanged. In either case the pattern would have no power, as it were, to extend itself. Or to state it yet another way, the pattern would not by itself constitute a rule that the coin or die will fall one way as opposed to another. Neither will the pattern amount to a rule that one outcome will now become more likely than another.

Two observations need to be made about the chance patterns I have just described. First, size and complexity are irrelevant to the question of a pattern's influence over future instances. It would be extremely unlikely that by tossing a fair coin we would generate a run of one hundred

heads, or by rolling a fair die generate a string of numbers that happened to fall in a sophisticated mathematical sequence. But in principle such large, chance patterns could occur, and if they did they still would not alter in the slightest the odds for the next coin toss or die roll.

The second observation has to do with the qualification I just made that we have in view a "fair" coin and die. To the extent that the coin and die were fair, our hypothetical patterns would owe to chance and nothing else. A chance pattern, in other words, is a pattern existing by itself. If the coin or die were "loaded" then the pattern would owe to a rule. For example, if a small, off-center weight were built into the die, then the resulting pattern of rolls would be due in part at least to the rule known as the law of gravity, or properly, the Law of Universal Gravitation. The rule that generated the pattern, rather than the pattern as such, would be the source of predictability about the next die roll.

To repeat, a pattern by itself is not a rule and offers us no predictive insight. In the case of the loaded die we have something more than a pattern, namely a rule in the form of a physical law. Or do we? Hume would point out that what we call the Law of Universal Gravitation is really just a pattern humans have observed in nature. We may try to take refuge in the immense size and orderliness of this pattern. Humans have made billions of observations of objects falling and of planets orbiting, and the pattern of such observations irresistibly suggests a rule. However, we just agreed above that the size and statistical improbability of a pattern are irrelevant to the question of whether "by itself" it has the force of a rule--it does not. We thought that in the case of a loaded die we had something more than a pattern to offer us predictive insight, but now we realize that the something more is--just another pattern.

We might escape this dilemma if we admitted that the rule known as the Law of Universal Gravitation is something other than the pattern it generates in the physical universe. That is a step that most naturalists, and particularly physicalists, are loath to take. Look at the way Richard Carrier explains physical laws:

> Thus, all you need for the laws to exist is a universe that works a certain way--it will automatically follow from the existence of any such universe that there will be a best way to describe and manipulate it, and that "best way" is something humans can discover and then describe. We then name these descriptions "rules" or "laws."

The "laws of physics," Carrier claims, are simply "descriptions" of the way the universe "actually, physically behaves," or to be more precise, of the way human beings have until now observed the universe to behave. A physicist might additionally refer to "regularities" to highlight a consistency in nature that makes it intelligible. This amounts to a claim that some patterns are rules. Patterns are never rules, as we saw in the case of die rolls. A pattern existing by itself or on its own is indistinguishable from a pattern arising by chance, and a chance pattern offers no predictive insight.

Elsewhere Carrier says that "the law of gravity exists as a fact of the universe as a whole--for there is no particular 'place' or 'time' where the law of gravity 'is,' it 'is' in every place and time where the physical conditions that manifest gravity exist." There is a difference between saying that the law of gravity exists at no particular time and place and saying that it exists at all the times and places where gravity is observed. Gravity is one thing, the law of gravity another. To identify the law of gravity with instances of gravity is a classic confusion of rule with pattern.

The obstinacy of the problem of induction is well-known within the philosophy of science. If we don't hear more about it, it is because philosophers tire of rehashing conceptual difficulties for which no solution is to be found within the bounds of naturalism. Scientists in general need

not grapple with this puzzle in order to carry on their research. Einstein apparently gave it some thought, however, because he acknowledged that the existence of rules in the form of physical laws cannot be proven[13]. When we give up the assumption that mindless matter is the bedrock of reality, we find that there is a way forward.

Since we cannot identify rules as the patterns they generate, we are left with two possibilities. The first is that rules are mere illusion, as is our ability to predict the behavior of physical objects such as falling apples and orbiting planets. The second, infinitely more attractive alternative, is that rules exist independently of the patterns they generate. But how can rules exist apart from their physical manifestations? We have a clue in the fact that rules, including physical laws, are propositions. The Law of Universal Gravitation, the laws of electromagnetism, and the formulations of quantum mechanics, among others, all have the form of equations to which variables can be assigned to create prose sentences. As should be obvious, it is the meanings of these sentences and not their physical representations that condition the physical universe[14]. Fresh from our discussion of the argument from reason, this is familiar territory. By invoking the meanings of propositions, we begin to envision a foundational role for things other than material objects and states.

Our second clue about the nature of rules is the fact that human minds can invent them. Not only do we invent rules, the rules we invent generate physical patterns. Take, for example, the rules of chess. These rules result in physical patterns in the movements of the pieces on a chessboard. But where do the rules exist? If we say that they exist in books about chess, we will have to correct ourselves immediately. Books contain representations of the rules of chess, but not the rules themselves. Do the rules exist in human brains? Finding a place for rules in the brain, considered as a purely physical system, is impossible for two reasons. For one thing, rules take the form of propositions, and as we saw under the topic of mental causation, the states of physical structures like the brain can qualify as representations of propositions, but not as propositions themselves. For another, the neural firing sequences that accompany chess play constitute patterns and, as we just considered, patterns cannot be identified with rules.

These facts notwithstanding, the rules of chess are real and play an undeniable causal role in the physical world: chess players directly experience that they move the pieces according to the rules of chess. If we deny rules this role, we risk our grasp on the world as an intelligible place. Humans create rules, not just for playing games, but also for conducting and interpreting scientific experiments. Moreover, a claim that rules are illusory falls prey to the kind of coherence argument we have already considered above regarding volition. Therefore, even if we cannot picture how rules influence the course of physical reality, we are forced to give them their due.

A clarification is in order as to the type of rules we are now considering. We are talking about rules that lack the force of logical necessity. Rules that constitute necessary truths, such as the principles of logic and mathematics, do not fall under the problem of induction and are not inferred from patterns. We do not conclude that one plus one invariably equals two because we have seen over and over again that they add up that way, but because it is impossible to conceive of them adding up otherwise, in this or any universe. Necessary truths are understood or perhaps discovered, but not invented. We cannot imagine that they began to obtain at some time in the past or may cease to obtain in the future.

Let me emphasize here that the laws of physics, like laws of nature generally, do not have the force of logical necessity. In that respect the laws of physics are more like the rules of chess than they are like the principles of mathematics. We can imagine the rules of chess being different than they are. Likewise, we can imagine the law of universal gravitation operating with a different mathematical value or even ceasing to obtain in the future. What does this tell us? The reason that we can invent alternate rules of chess is because the current rules were themselves invented. The rules of chess make a difference in the physical world through the

exercise of human will, forming patterns in the activity of the brain, nerves, muscles, and finally chess pieces. Only a much greater mind and will than that of human beings could cause the laws of nature to generate patterns in the activity of all matter and energy.[15]

A different way to appreciate this point is by considering alternate universes. The set of all logically possible universes includes some universes in which the laws of nature change at certain times as a matter of brute fact, without further explanation. In one of these universes, for example, the value of acceleration due to gravity changes without further explanation on August 1, 2010. Such a universe is observationally indistinguishable from our own prior to that date. So is any universe in which the mathematical parameters of the laws of nature remain constant during a time envelope that, by chance, overlaps the time envelope of human observation. However, in all such universes the apparently nonaccidental constancy of physical laws over time is illusory because the overlap of envelopes is accidental.

As an illustration, imagine that you stand in front of a gift-wrapped package. Behind it, other gift-wrapped packages stretch out to infinity. Some of these packages are empty and some contain gifts. If you had a choice, you would deliberately stand in front of a package containing a gift. You do not have that choice, but nevertheless you are intuitively convinced that the package in front of you contains a gift, and that it does so nonaccidentally. How can your intuition be correct if the only way you could stand in front of a gift is by chance? Your package contains a gift nonaccidentally only if it does so as the result of a purposeful act comparable to your own hypothetical act of choosing to stand in front of a package that you know contains a gift. In order to validate your intuition, there must exist a rational agent who, unlike yourself, has the power either to place a gift in your package or otherwise arrange for you to stand in front of a package containing a gift. In our case, the "gift" is the intelligible consistency of the rules of the universe over time and across space.

Some may argue that all rules--even the rules of chess--have always existed in a timeless, spaceless Platonic realm. But in contrast to logical and mathematical laws, the rules of chess appear to have come into existence and thereafter played a role in the world when human minds invented them. And the power of those rules to affect physical reality derives completely from human minds. To say that these rules exist "in" our minds is a metaphor not necessarily better than saying that they exist "to" our minds or "for" them. But the need for the metaphor does not alter the reality of the rules of chess, their origin, and what their origin tells us about the laws of nature. We have seen that like man-made rules, the laws of nature lack the force of logical necessity and must therefore draw their power from thought and volition. However, the laws of the universe have primary status, while man-made rules are secondary. The rules of chess, traffic laws, and building codes are layered on top of, and depend upon, these primary rules of nature. This difference between the types of rules must reflect a difference in the originating minds in either case. The mind that generates the primary rules must be primary by comparison with our own dependent and derivative minds. And precisely because our own minds but faintly echo the transcendent mind, we cannot understand in all respects its mode of existence.

The fact that mental activity must underlie the laws of nature does not by itself tell us whether we should credit *a* mind, or minds plural. If a single mind were responsible, we would have to call it God. That we cannot see God's thoughts conditioning the course of the universe is not surprising, since we cannot picture the way rules in our own minds influence the course of events in our brains, bodies, and surroundings.

In our experience a mind always requires a brain, but the mind that confers intelligibility upon the universe would have to transcend bodily existence such as we have. Admittedly, this is a challenging idea. But we have just considered a powerful inference in its favor.[16] And by demonstrating that even our own rationality defies containment within physical processes, the AfR lends support to the notion of a transcendent mind.

Where did the transcendent mind come from? If the intelligibility of the universe owes to this mind, then the universe as an intelligible construct cannot have preceded it. The transcendent mind is necessarily eternal, or at least not time-bound in the sense that physical states are, considering that it provides the context necessary for physical states to undergo change.

I referred originally to a mind or minds, but a single mind is the simplest explanation. There is another reason for thinking that one mind is more likely. While human conceptions and constructions can be created cooperatively, the most evocative of these are works of art that originate in the unified vision of one mind. If the universe is more like a work of art than a machine, or if though being a machine it is at the same time a sublime work of art, then a single mind is likely responsible. I don't pretend that this is other than a suggestive argument, but it still carries some weight.

We cannot speculate about a transcendent, universe-governing mind without thinking of the world's religious traditions and their claims of revelation about God, gods, or some greater reality beyond the reach of human senses. If we concede that possessing a mind is the essence of personhood, then we have not come out far from the Judeo-Christian image of a transcendent universal Creator who by means of cosmic ordinances "leads forth" the constellations, "forms the mountains, creates the wind and reveals his thoughts to man" (Job 38:31-33; Amos 4:13).

The idea that God is the conditioning intelligence behind the universe may be shown more or less likely through logical argument, but it cannot be tested as a scientific hypothesis. Does that fact support the secular charge that God has no explanatory value? Consider the statement, "Nature conforms to laws." It may surprise some readers to learn that this statement is not a scientific theory or even a hypothesis. I already noted that Einstein said, correctly, that the conformity of nature to rules cannot be proven. To qualify as a hypothesis such conformity would have to be capable of being falsified, which clearly it is not. It seems then that all that is explanatory is not strictly scientific, even though it may bear upon science. If nature's conformity to laws points logically to a divine mind, then that idea has explanatory worth.

Morals and ethics provide another example of explanation apart from scientific testability. Secularists deny that belief in God is necessary for morality, but I doubt that most thoughtful secularists deny the reality of objective moral values as such. Can science demonstrate for us what these values are or even that they exist? Surely Bertrand Russell was right in saying that it cannot[17]. What experiment could prove that maliciousness and treachery are morally wrong, or that generosity and honesty are morally right? Even if moral values have no scientifically testable implications, most of us credit them with explanatory value in the case of our own behavior. Few of us would deny that they are worth knowing about. God's existence is also worth knowing about for spiritual and moral reasons, whether or not it is grist for scientific inquiry.

What about a God who performs miracles? Are we bound to reject violations of nature's laws on grounds of science? Go back to the rules of chess. In chess there is a maneuver called "castling" in which under special circumstances the king and rook move in ways that they ordinarily cannot. Castling would momentarily confuse someone attempting to decipher the rules of chess based on patterns in the movement of pieces. Castling is not a violation of the rules of chess, but it does show that those rules are more complicated than they at first appear. The complication represented by castling only occurs under special, infrequent circumstances. Biblical miracles such as the parting of the Red Sea or the resurrection of Jesus have a spiritual rationale that provides for the relentless quality of nature's laws under all but certain rare (and not necessarily recurring) historical circumstances. The reason that they can occur as complications of the pattern of nature is the same reason that nature proceeds with strict regularity the rest of the time. Given the argument from intelligibility, miracles are, in the words of the cliché, exceptions that prove the rule[18].

A pair of disclaimers: The argument from intelligibility is not a formulaic "proof" of God. I have outlined a problem and then referred to clues that point toward the solution lying in one direction rather than another. That this direction is compatible with Christian theism is significant, but by itself is unlikely to be decisive for anyone.

Similarly, the AfR does not constitute an airtight theistic proof. Still less is it proof that each of us has a soul in the form of a ghost hovering inside our skulls, as C.S. Lewis himself was quick to observe. To Lewis, the AfR showed that when we reason something occurs that is impossible to capture in a natural (mechanistic) account of brain function. A two-fold dependence must therefore exist. Our dependence on the physical organization of the brain is accompanied by a dependence on God's all-conditioning thought, which alone can liberate cognitive processes from the straightjacket of physical cause-and-effect. Lewis's account coincides with the Christian claim that God is the one in whom, even if heedlessly, "we live and move and have our being" (Acts 17:28). This is especially so since, if the argument from intelligibility has substance, even brain function in purely physical respects depends on the background rationality provided by God. Disclaimers aside, the biblical prospect of life beyond death looks more probable if the AfR has merit.

Integrating the Arguments from Reason and Intelligibility

The arguments I have attempted to clarify here are interrelated. Naturalism implies that the universe is a brute fact. By "universe" I refer not just to the state of matter, energy, and space at a certain time, but all states and events that have occurred or ever will occur. Naturalism ultimately dictates that the universe is as it is because it happens to be that way, not because an originating mind intelligently conceived it. The dice on the tabletop fall the way that they do because the molecular dice fall the way that they do. Various layers of explanation are distinguishable only by the shape of the dice being rolled at each level, never by the brute fact quality of the explanation.

However, if brute factuality were the basic property of the universe, it would be the basic property of our minds as well. The events in our minds, being part of this world, would share its accidental character. Whatever our illusions about rational relationships between ideas, the following of one thought by another would be an undistinguished droplet in the bottomless sea of brute facts.

Such an account is fundamentally at odds with rationality as we experience it. A sequence of thought cannot just happen to be logical. The word "brute," as in brute fact, has among its meanings "lacking rationality." If we say that logic itself just happens to be the way that it is, we turn out to be saying something different than when we say that events just happen to occur in the order they do. When we describe a sequence of thought as logical, we add an irreducible layer of explanation even if someone presses the claim that logic in some sense happens to be the way that it is.

As it is with rationality, so it is with volition. To say that an act is intentional (as opposed to unintentional) is to add an irreducible quality to that act. We can say that an act just happens to be intentional, but if this claim were to reduce the act to the status of brute fact it would imply that we don't know what we mean when we talk about an act being intended versus unintended. But we do know what we mean. To acknowledge--as we must--the irreducible phenomena of reason and volition in our own case, we must abandon the brute fact account of the universe to which we belong.

We cannot deny that reason and volition appear only when a complex and delicately balanced physical structure, the brain, is operating. The brain must be like an automated telescope, full of complex gadgetry that allows it to track with precision a point source of radiance far out in the

cosmos. As long as the device functions properly, the radiant source floods the internals of the machine with something we might call, for lack of a better term, enlightenment. This nonphysical, psychic energy allow events to occur that would not otherwise, consisting of rational insights and intentional acts. If the gadgetry is interfered with, the supply of this enlightenment is disrupted and these types of events cease to occur[19].

Therefore, matter, energy, and space are not the fundamental constituents of all else. Rational thought is not an arrangement of matter or series of physical states. We will never understand rationality in terms of particles thoughtlessly going about their business. Whoever believed that we might better understand logical relations between propositions by studying brain chemistry can now be disabused of the error. But where do we go from here?

One of the uses to which we put intention and reason is the creation of virtual models, as Richard Carrier claims. Among these is a virtual model of the universe itself, driven by scientific laws that we apply to it. We must assume that our mental model of the cosmos, whatever its imperfections, is intelligently conceived. But how can we hope that this model corresponds to the actual universe unless the real article has been similarly conceived? If the intelligibility, not to mention accuracy, of our internal model of the world owes to the exercise of our creative intelligence, how can the archetype that it aspires to mimic possess that same quality apart from the exercise of a creative intelligence deeper and vaster than our own? And doesn't our own near-obsessive determination to engage in virtual model construction bear witness that we are made in the image of the one who conceived the archetype?

In his *Dialogues Concerning Natural Religion*, Hume made the grudged admission (placed on the lips of his character Philo) that the "cause or causes" of order in the universe "probably bear some remote analogy to human intelligence."[20] Hume had already qualified the statement by commenting that the world contains so much suffering that we can hardly believe it to be the creation of a caring God. Yet by admitting that the universe seems to draw its coherence from an intelligent cause (or causes), Hume revealed that the argument from suffering had been the mainstay of his religious skepticism all along. Hume, in other words, implies a difference between the beliefs that (1) behind the universe is an all-governing intelligence and (2) behind the universe is an intelligence *that cares about us*. I believe that both propositions are true, although the question of whether God can be a loving being is beyond the scope of this paper. But even if we only get as far as a universe-governing mind we have made progress.

I would even say that we have made progress if fewer of us assume that the arguments from reason and intelligibility are soft targets.

I have tried here to convey the AfR in terms somewhat different from those in Victor Reppert's book. The reason is not because *CSLDI* is deficient, but because in light of Richard Carrier's critique a fresh approach seemed appropriate. If what I have said here changes few minds, perhaps nevertheless it will help to move the discussion forward.

Notes

[1] I wish to acknowledge the contribution of Keith Augustine, whose incisive questions, comments, and editorial suggestions were of great help in the writing of this essay.

[2] For a balanced discussion, including literature citations, see Alex Byrne's entry on "Intentionality" from *Philosophy of Science: An Encyclopedia*, ed. J. Pfeifer and S. Sarkar (Routledge, forthcoming), <http://mit.edu/abyrne/www/intentionality.html>.

[3] Though we are concerned at this point only with perception of logical relations, philosophers have claimed that a number of other phenomena are available only to introspection and may therefore pose problems for naturalism.

[4] Supervenience theory cannot help with this problem because propositions do not supervene on their representations. Supervenience requires covariance out of logical or metaphysical necessity, a much stronger relation than that of representation.

[5] Carrier says: "If a machine can perform acts of rational decision (and machines certainly can, both deductive and inductive), and if another machine can 'perceive' that activity and construct a corresponding virtual model of it (and machines can certainly do this, too), then where is Hasker's violation of causal closure?"

[6] Richard Nisbett and Timothy Wilson, "Telling More than We Can Know: Verbal Reports on Mental Processes," *Psychological Review* 84(3): 231-259 (May 1977).

[7] Ibid., p. 247.

[8] For a brief argument for epiphenomenalism from a leading researcher in the fields of human cognition and artificial intelligence, see Stevan Harnad, "Turing Indistinguishability and the Blind Watchmaker" in *Consciousness Evolving*, ed. J. Fetzer (2002), p. 4. See also by the same author various online articles on consciousness, behavior, and symbol-grounding.

[9] See David Hume, *A Treatise of Human Nature*, 1.1.

[10] Bertrand Russell defined the paradoxical set as the set of all sets that do not contain themselves as members. The popular form of the paradox describes a village with the rule that all men must choose between either shaving themselves or else being shaved by the local barber. How can the barber himself abide by the rule, assuming the barber is a man?

[11] See John Foster, *The Divine Lawmaker: Lectures on Induction, Laws of Nature and the Existence of God* (Oxford University Press, 2004), pp. 17-35.

[12] See, for example, Alexander Rosenberg, *Philosophy of Science* (Routledge, 2000), pp. 33ff.

[13] "It is the aim of science to establish general rules which determine the reciprocal connection of objects and events in time and space. For these rules, or laws of nature, absolutely general validity is required--not proven." Albert Einstein, *Ideas and Opinions* (Bonanza, 1954), p. 47.

[14] This corresponds to our previous observation that representations of propositions cannot be given the causal role of propositions themselves in reasoning. Representations of rules cannot coherently be said to serve the causal role of rules, as they must be if the mind is nothing more than the physical system of the brain.

[15] Throughout this essay, I use "introspective" and "introspection" broadly to identify perception that is nonsensory. Besides logical relations and abstractions, philosophers generally recognize other entities that are known apart from direct or indirect sensory contact, such as memories. More controversial as possible objects of nonsensory contact are, for example, moral facts and mystical states.

[16] For a more developed treatment of the subject see Hugo Meynell, *The Intelligible Universe: A Cosmological Argument* (Barnes & Noble, 1982). See also by the same author: "Hume, Kant and Rational Theism," *Truth Journal*, Vol. 3 (1991), <http://www.leaderu.com/truth/3truth08.html>.

[17] "There remains, however, a vast field, traditionally included in philosophy, where scientific methods are inadequate. This field includes ultimate questions of value; science alone, for example, cannot prove that it is bad to enjoy the infliction of cruelty. Whatever can be known can be known by means of science; but things which are legitimately matters of feeling lie outside its province." Bertrand Russell, *A History of Western Philosophy* (1945), p. 834.

[18] See the discussion by Victor Reppert, "Hume on Miracles, Frequencies, and Prior Probabilities," <http://www.infidels.org/library/modern/victor_reppert/miracles.html>.

[19] Since I am not referring to energy in the physical sense, this image is only a metaphor. However, even in science there are concepts that cannot be pictured in the mind and can be grasped only by metaphor. The general theory of relativity is commonly represented as a bowling ball resting on a mattress; the expansion of the universe is pictured by patches on an inflating balloon; the existence of more than three spatial dimensions is diagramed in various ways but cannot be visualized; etc.

[20] David Hume, *Dialogues Concerning Natural Religion*, p. 227 in the Norman Kemp Smith edition.
