

Can Philosophy of Mind Provide Reasons for Believing in God?

by Michael J. Degnan
Department of Philosophy

While it is true that certain religious beliefs are at odds with well-confirmed scientific theories, most theologians respond that such beliefs are not necessary to many mainline Christian, Jewish or Islamic religions.¹ However, some mainline Christian interpretations of human freedom do appear inconsistent with the physical determinism allegedly assumed by the natural sciences. The Roman Catholic tradition, for example, insists on a robust sense of free will. That tradition affirms that human agents have the ability to do otherwise than they in fact do given no change in the initial conditions. But this strong view of free will seems to entail the falsity of the determinism allegedly assumed in physics and chemistry.

In this paper I will argue that a proper understanding of the determinism characteristic of physics and chemistry does not require the denial of the robust sense of human free will. Philosophy of science can reconcile what appears to be an inconsistency between the strong view of free will and the determinism in natural science. Because many philosophers of mind believe that the integrity of the natural sciences demands the falsity of human free will, they conclude that physicalistic accounts of human thinking must be correct. If the existence of human free will does not have this consequence, there may be reasons for preferring nonphysicalistic accounts of thinking. In the second part of the paper I will argue that features of consciousness and thinking that two prominent Anglo-American philosophers use to support a nuanced physicalist account of thinking actually point to the necessity of a supernatural intelligence.

1. Are Free Will and the Determinism of Science Incompatible?

In *Minds, Brains and Science* John Searle, a prominent American philosopher, argues that all mental phenomena are simply features of the brain because all mental phenomena are caused by processes going on in the brain. From this he concludes that humans have no free will. His argument can be

A. Any system that is simply a feature of physical and chemical laws is a deterministic system (“simply a feature” means only a feature).

B. The human mind is a system that is simply a feature of physical and chemical laws.

C. So the human mind is a deterministic system.

D. No deterministic system is a system that exhibits free will.

E. So no human mind exhibits free will.

Despite this argument Searle concedes that he cannot abandon the belief that he can do otherwise than he in fact does. Searle believes that this belief is genetically programmed into him. According to Searle having and acting on this false belief has some kind of survival advantage for the human species. The point of interest for this paper is that Searle believes that the truth of the free will thesis is inconsistent with the truth of a key assumption undergirding the natural sciences. Searle's view of this incompatibility is prevalent among contemporary analytic philosophers.²

I will now argue that the argument A–E above is flawed because premise B is false. While it is true that any system that is simply a feature of physical-chemical laws is deterministic, it does not follow from this that every concrete realization of such a system is deterministic. To think that the determinism characteristic of the system in abstraction from the concrete applies to the concrete is an example of what Alfred North Whitehead calls, “the fallacy of misplaced concreteness.”³ In order to explain this fallacy it is helpful to consider the relation between concrete events and scientific theories. We can characterize the relation in terms of

1. A law, statable in principle in terms of the primitives of the system, which asserts that if the relevant traits of the object(s) are in one state at one time they will be in another state at another time, or that if some relevant traits are in one state other traits will be in another state. I shall refer to the relevant traits as “state variables.”
2. Statements of initial conditions noting what is the case in the actual phenomena with respect to the state-variables used in the antecedent of the law.
3. Closure assumptions which are suppositions that traits of concrete objects that are not referred to in the law do not ever or do not always affect the interactions of traits that are referred to in the law.

(Homogenous closure assumptions are suppositions that the interrelations of relevant traits in one concrete object or system of objects can be adequately described without reference to the same traits of other concrete objects. Heterogenous closure assumptions are suppositions that the traits selected in the abstraction process are functionally independent of those traits of the concrete situation not selected. For example, in a study of the Galilean equations of free fall motion of objects in introductory physics classes, the instructor asks the students to neglect the factor of air resistance, for it is a trait that belongs to the concrete object that is deemed not important when one is dropping lead weights off of skyscrapers).

A vivid example of the interplay of these factors is given by Harry Austin:

Being rather strait-laced and unimaginative, I am spending my holiday in Paris amusing myself by flipping small stones off the Eiffel tower and trying to predict their impact points. Using the axioms of Newtonian mechanics, the law of gravity and the initial

conditions of the stones stated in terms of the mechanical primitives of the system, position and momentum let us say, I am quite successful except for a small group of stones which always seems to deviate slightly from predicted impact points in the direction of the tower. Being quite full of non-Parisian ideas, it occurs to me that this deviant subgroup might be comprised of lodestones, which, no doubt, would be deflected in their paths by the iron in the tower. Taking them to my hotel room I test them with my vade-mecum galvanometer circuit, and get a needle deflection. I am now satisfied that if I bone up on the laws of magnetism and make certain initial measurements, I shall be able, armed with two sets of laws in two vocabularies as well as with two sets of initial conditions, to predict fairly accurately the impact points of the deviant stones.⁵

There are two items to notice about this example. First the hypothesis that the deviant stones were magnetized had to be confirmed independently of the fact that they deviated from predicted impact points. Second, the actual courses of some of the stones did not conform to the laws of Newtonian mechanics because the heterogenous closure assumption did not hold. The lodestones possess a trait that is neglected by Newtonian mechanics, the magnetic vector. As long as heterogenous closure assumptions hold, the determinism that holds at the level of the abstractions also holds at the level of the concrete. But when there is a breakdown in the closure assumptions the determinism that holds at the level of the abstractions is not guaranteed in the concrete case. If there were no concrete objects with respect to which Newtonian mechanics were almost completely deterministic, the theory could receive no confirmation. But just because the assumptions hold in some cases is not reason to think that they always hold. On the other hand, the breakdown of closure does not imply that the laws of mechanics do not obtain in such situations, but rather what applies is not only the laws of mechanics. When we apply this material to the explanation of human behavior, say the explanation of one of you getting up to get a glass of punch at the back of the room, we see that it is possible that the physical-chemical explanation of your movement will fail, not because the physical-chemical laws do not ever apply to concrete cases, but because what applies is not only such laws. If the existence of subjective factors can be independently established, as in the lodestone case with the galvanometer, laws other than those of physics might be needed to explain the fact that one of you was moving toward the punch bowl. However, unlike the lodestone case where the other factor could be added to the factors already mentioned, the subjective factor is not so well-formulated. If this subjective factor cannot be expressed in law-like formulas, then there is room for free will. This means that one can affirm free will without holding that the laws of physics and chemistry are false. Rather, one can affirm that the systems for which the heterogenous closure assumptions hold are indeed deterministic in concrete instances. The success of science is explained. But the inability of some social sciences to succeed in giving explanations as general as those given in physics and chemistry is explained as well, for often the very traits left out of the primitives of a social theory are just the ones that make all the difference when applied generally. This was surely the lesson for many macro-economists who continued to be amazed at the strength of the economic recovery during the Reagan years. Factors like motivation of the entrepreneur were not considered in their models. This analysis of the relation of scientific theories to the concrete does not prove that there is free will. It simply makes conceptual room for attributing free will to human agents, for it shows that

we can affirm that scientific theories apply to concrete cases just as long as closure assumptions hold. But if these assumptions fail when it comes to explaining human behavior, it may be because of a subjective factor like free will. This account removes an important reason for accepting the claim that the human mind is a system that is only a feature of physical and chemical laws. The human mind that causes us to act and is allegedly the source of free will is not a system. It is a concrete thing. Just because determinism holds at the level of abstraction, it does not follow that it also holds at the concrete level. One of the reasons Searle endorses the claim is the fear that if he allows there to be some feature of the mind not caused by the brain then there appears to be a breakdown in the laws of physics. If that happens, then all hell breaks loose. So rather than give up on science he relinquishes free will. On the contrary, I have argued that all hell will not break loose. One simply recognizes that with mental phenomena closure assumptions break down. There is conceptual room for theological or philosophical explanations of human behavior that leaves untouched the important successes of the natural sciences.

2. Consciousness as a Radically Emergent Feature

If there is conceptual room for critiquing the adequacy of the scientific account in the way I have described, we need to see what kinds of arguments might fit the bill. Aristotle in the *De Anima*, a study of the life principle in organisms, and Aquinas in his commentary on that work, both argue to the existence of a supernatural intelligence on the basis of their analysis of human thinking.⁶ They reason that if there are kinds of entities whose properties cannot be explained unless we infer the existence of a supernatural intelligence, then we will have important evidence for God's existence. Human beings are such entities since there is no adequate naturalistic explanation of thinking and consciousness. By contrast, most twentieth century Anglo-American philosophers do not take this lack of explanation as having such momentous consequences. Some, like John Searle in *Rediscovery of the Mind*, argue that this lack is temporary. It is just a matter of time before neurobiology gives us a successful explanation of human thinking and consciousness without any reference to nonphysical intelligences. Searle argues for this possibility even while recognizing the irreducibly subjective aspect of consciousness and the causal power of mental states. On the other hand, Colin McGinn in *Problems of Consciousness* argues that this lack is necessary given the kind of understanding we humans can achieve. The lack of explanation does not entail the existence of a supernatural being, for his thesis is about the limitations of human knowledge, not the character of consciousness per se. Just as the fact that armadillos cannot solve problems of elementary arithmetic does not entail anything mysterious about arithmetic, so the fact that humans cannot solve the problem of consciousness does not entail anything mysterious about consciousness. Like Searle, McGinn believes that consciousness is caused by some combination of physical and chemical interactions in the brain. We just cannot explain how such interactions issue in all the wonders associated with thinking and consciousness. Both Searle and McGinn agree that consciousness is not identical with some discrete physical part of a human being. They see it as a system feature, to use Terry Nichols's phrase, they see it as a holistic property. However, not all holistic properties entail the falsity of a wholly

materialistic explanation. McGinn and Searle understand consciousness to be a “causally emergent system feature” which is a result of the interaction among many physical processes going on in a human being.⁷ Consciousness is explained as a consequence of the interaction of many processes just as the liquidity of water is explained as a consequence of the interaction of the atomic bonds between hydrogen and oxygen atoms. Hence the existence of consciousness in humans does not require any inference to a supernatural being. They both deny that consciousness is a “radically emergent system feature” another kind of holistic property. A radically emergent system feature is one which has causal powers that are not fully caused by the interactions of the elements of the system, but the interaction of the elements is necessary for the existence of the system feature. An intrinsic principle of organization characteristic of that type of whole is responsible for this type of emergent system property. This kind of holistic property may entail the existence of a supernatural being. In the next sections of the paper I will argue that some of the beliefs Searle and McGinn use to support the causally emergent characterization of consciousness entail that consciousness is a radically emergent system feature. If I am correct, then the features of consciousness and thinking they point to have consequences for the existence of a supernatural being. Before I can criticize them in this way I must consider the rejection of the very possibility of radically emergent system features.

3. An In-Principle Argument Against Radically Emergent Features

Searle argues against the possibility of any radically emergent system features, which he calls emergent² properties. His argument can be expressed as follows:⁸

1. If there are emergent² properties, then the principle of transitivity of causation is false.
2. The principle of transitivity is true.
3. There are not emergent² properties.

I understand the principle of transitivity of causation to be the following claim: If A, B and C are entities or events and if A causes B and B causes C, then A causes C. There are many causal relations that confirm the truth of this principle. Whenever A is an agent stimulating some other instrumental agent to act, the principle is true. For example, if I hit the cue ball and my hitting causes the movement of the cue ball which causes the movement of the eight ball, then my hitting the cue ball caused the movement of the eight ball. But the principle has more support than is given by lots of examples, for the very intelligibility of thinking about causal relations seems to require its truth. If we deny this principle, then we must modify our notions of causality as being a kind of dependency relation. Although Searle is not explicit about this, the principle seems to be a necessary truth or a first principle in the way Aristotle uses the phrase. The case for the truth of premise 1 also seems strong. Assume we have an entity composed of a number of elements. Assume also that we know the causal series that results in interaction C which is the cause of an emergent feature E of the entity in question. C is caused by B which is caused by A. Emergent feature E is causally responsible for feature D. According to the

principle of transitivity, A is the cause of D. But such is not the case if E is an emergent² property since the causal power of an emergent² property is not a function of the interrelation of the parts. Hence, if there are emergent² properties, then the transitivity principle is false. Searle's argument is very important, I think, because it expresses the viewpoint of many workers in the natural sciences with whom I have talked about these matters. Last spring, for example, I invited Dr. Virgil Mathiowetz, an internationally recognized worker in occupational therapy, to speak to my human person class about a model of the human being he and his coworkers had developed for therapists to use. The model views the human being as a complex system of interacting principles some of which are biological, some psychological, some economic, some religious and so on. I asked him whether he thought in principle that he could explain the features of this whole system once he understood the interaction of the parts. He agreed that he could, although he admitted that in fact he had no such explanation. When I asked if he thought that the features could not be explained by the interaction of the parts he seemed puzzled. He told me that he thought such a state of affairs would mean that his discipline had failed at explaining the phenomena. The very paradigm of explanation that he used would not fit the data. Similarly, in *Towards a New Philosophy of Biology* Ernst Mayr argues that an organism is an integrated whole and not a collection of independent genes.⁹ But he takes the whole as something that can be explained in terms of the interaction of the parts and processes. To deny this seems tantamount to failure of the discipline. So Searle's argument succinctly expresses the in-principle worry emergent² properties present for workers in the natural sciences.

4. Response to Searle's In-Principle Argument

I will argue that premise 1 of Searle's in-principle argument is false because it is possible that an emergent² feature exists without the falsity of the transitivity principle. In the reasoning given above it is assumed that E, the emergent feature, is caused by the interaction C. But that assumption cannot be the case, if E is an emergent² feature. The interaction C may well be a necessary condition for E, but it is not sufficient because the new causal power in E is not caused by C. There must be something additional to C and its antecedents that causes this new causal power. The reason there must be something additional is based on the claim: The effect of any given cause cannot exceed the causal power of that effect's cause.¹⁰ For example, a pot of water on the stove cannot be heated to a temperature above the temperature of the flame that is heating it. This is analogous to the second law of thermodynamics. Causal power is likened to energy. If the causal power is not caused by the interaction of the elements, then there must be some other causal force involved than the interaction. It is this other causal force in addition to the interaction that causes the emergent feature. The claim that the effect of any given cause cannot exceed the causal power of that effect's cause needs to be supported by more than a casual example about heating pots of water. First, the principle is about one kind of cause, active or efficient causes. An active cause is one that brings about things that come to be. Heat is an active cause for it brings about the change from solid to a liquid, for example. Gravity is another active cause responsible for the formation of stars, planets and galaxies. Because the pressure within a star results in an increase in temperature, and since the pressure within a star is caused by gravity, by the transitive principle, gravity is

the active cause of the rise in temperature. A consequence of this causal principle is that it is not possible for something to come to be from nothing. But a theist might say that this is just what happens in creation *ex nihilo*, from nothing God creates something. There is no need of a material cause in this coming to be, i.e. no need of some preexisting material or subject for the coming to be. But there is a need of an agent, in this case God, who in some fashion actually contains the effect to be caused. The claim at issue is “The effect of any given active cause cannot exceed the causal power of that active cause”. In other words, every active cause acts in the measure of its actuality and cannot act beyond that actuality. The actuality of an active cause includes the power and capacities which exist here and now. When the meaning of this claim is spelled out in this manner the claim is seen to be true by definition. Hence, I offer this claim as a necessary truth. If this causal principle is true, then I have shown that the existence of emergent² properties does not entail the falsity of the transitivity principle (however, I have not shown that there are any emergent² properties yet). If “cause” simply signifies a necessary condition, then it is true that the necessary conditions of the interactions antecedent to the coming to be of the emergent feature are also necessary for the emergent feature and the emergent feature’s causal power. However, if “cause” signifies a necessary and sufficient condition, then the interactions antecedent to the coming to be of the emergent feature are not the cause of that feature. If we include this other causal force in addition to the interactions, then we have necessary and sufficient conditions. In that case, the cause will be both the cause of the emergent feature and the cause of the emergent feature’s causal power. I have argued that it is possible for emergent² properties to be real without entailing the violation of the transitivity principle.

5. Consciousness as an Emergent² Property

If I have succeeded in responding to the in-principle objections to emergent² properties, I am now ready to make a case for the reality of such properties based on claims that Searle and McGinn make in their books which allegedly have no room for such notions. Searle accepts the irreducibility of the subjective aspect of consciousness. He believes that a conscious organism’s experience of pain has a point of view. The experience is from the viewpoint of the experiencing organism. There is no way that we humans can know what it is like for a cat as a cat to experience pain.¹¹ There is good reason to think that there is a reality to such a subjective state, for we know that there is something it is like to experience pain for a human as a human. Thomas Nagel has argued that if an alien race came to earth they could not know what it is like to be a human for a human.¹² But that epistemic limitation would not entail that there is no reality to our subjective conscious experience. While some have taken this irreducibility to be an emergent² feature, Searle does not. He is a causal reductionist who believes that the causal powers of consciousness are entirely explicable in terms of the causal powers of brain processes. He does not claim that every feature of consciousness is identical with some physical property of matter. For example, some objects are solid and this has causal consequences. Solid objects are impenetrable by other objects; they are resistant to pressure, etc. But these causal powers can be causally explained by the causal powers of vibratory movements of molecules in lattice structures. Searle notes that in the history of science successful causal reductions lead to ontological reductions in which objects of certain

types can be shown to consist in nothing but objects of other types. In such cases we simply redefine the expression that denotes the reduced phenomena in such a way that the phenomena can now be identified with their causes. For example, color terms were once defined in terms of the subjective experience of color perceivers. Today physicists can define “red” as photon emissions of 600 nanometers; instances of the color red simply are identified with photon emissions of 600 nanometers.¹³ The subjective experience of red has been cut off from the “real” color. The general principle seems to be that if we achieve a successful causal reduction of a property, then when we siphon off the subjective element of such a property we get an ontological reduction. With consciousness, however, the very thing we are attempting to explain is the subjective element, so we cannot siphon that off and get an ontological reduction. Searle avers that we cannot reduce the subjective, conscious, first-person sensation of pain to the objective, third-person patterns of neuron firings in the thalamus and other regions of a person’s brain. No third-person description will convey the first-person character of pain. But this has no deep ontological consequences for Searle, since this irreducibility is simply a function of our definitional practices.¹⁴ The reductionistic practice is not designed to capture first-person features for these realities are appearances. Searle writes,

Indeed, it is a general feature of such reductions that the phenomenon is defined in terms of the “reality” and not in terms of the “appearance.” But we can’t make that sort of appearance-reality distinction for consciousness because consciousness consists in the appearances themselves. Where appearance is concerned we cannot make the appearance-reality distinction because the appearance is the reality.¹⁵

So far so good. There just is an irreducibly subjective component of physical reality. There is no reason to classify consciousness as emergent², for although it is not fully explained by the interaction of brain processes, inasmuch as it is an appearance, it does not have causal powers that fail to be explained. However, Searle assigns consciousness a causal role. Searle argues that consciousness is not a computer program for no program gives meaning, programs only require syntactical operations. Consciousness gives meaning. Hence, for Searle consciousness has causal power.¹⁶ Now things do not look so good for Searle.¹⁷ How can the irreducibility of consciousness have no troubling ontological consequences when it has causal power? Searle concedes that consciousness is irreducible due to the irreducibility of the subjective feature of consciousness. So consciousness is not simply a function of the interactions of the parts of the elements of the brain. This irreducibility was innocuous as long as no causal roles were assigned to consciousness. But when Searle gives consciousness the power to make meanings, he gives it causal power. It seems as if consciousness so described is an emergent² property: it has the causal power of making meanings which have a subjective feature which is not a function of the interaction of the elements that constitute the whole of which it is a feature. Such is my case for the Searle’s implicit recognition of the reality of an emergent² feature of consciousness.

6. McGinn’s Implicit Affirmation of Emergent² Properties

Like Searle, McGinn affirms the irreducibility of consciousness to physical processes, and he believes that consciousness is a causally emergent feature that is caused by natural processes in the brain. But unlike Searle, McGinn argues that it is logically impossible for humans to understand the nature of consciousness. Intelligences which form all their empirical concepts on the basis of introspection and perception are constitutionally precluded from grasping the nature of consciousness. Human intelligence is such an intelligence.¹⁸ McGinn assumes that human consciousness must be grounded in some naturalistic property P that is a property of the human brain. We can never know what that cause is in any determinate way, only that it is wholly naturalistic. We have already seen how little support McGinn has for the claim that this property P must be wholly naturalistic. My aim here is to show that certain features he assigns consciousness imply that it is an emergent² property. First, I will show that McGinn is wrong about the epistemically closed nature of consciousness, for if he is correct about that claim it may be thought that I cannot infer anything about the character of consciousness from descriptions of its powers. The irreducible character of subjectivity has deeper implications for McGinn than it does for Searle. He uses this feature about subjectivity to mount an argument for the claim that we cannot give a true psychophysical theory of mental phenomena.¹⁹ McGinn argues that P, the naturalistic property that grounds consciousness, cannot be known by introspection since no examination of consciousness itself will give us the property of the brain in virtue of which we are conscious. Second he argues that P is not a perceptible property of the brain. He writes,

Basically, I think, it is because the senses are geared to representing a spatial world; they essentially present things in space with spatially defined properties. But it is precisely such properties that seem inherently incapable of resolving the mind-body problem: we cannot link consciousness to the brain in virtue of spatial properties of the brain. There the brain is, an object of perception, laid out in space, containing spatially distributed processes; but consciousness defies explanation in such terms. Consciousness does not seem made up out of smaller spatial processes; yet perception of the brain seems limited to revealing such processes.²⁰

McGinn believes that consciousness is not constituted of smaller spatial processes because the “what it is like to be” aspect of consciousness is not a reality with spatial properties. This is the irreducibility of subjectivity that Searle affirms. But for McGinn this irreducibility has significant epistemic consequences. However, our inability to perceive P is not sufficient to justify epistemic closure, since P could be known from an inference in the way that a variety of theoretical objects are known: e.g. electrons, quarks, strong and weak forces. McGinn anticipates this objection, for he argues that P cannot be gotten at through inference:

4. No phenomenon that fails to be in the data to be explained is a phenomenon that can be explained by inference.
5. Consciousness is a phenomenon that fails to be in the data to be explained.
6. So consciousness is not a phenomenon that can be explained by inference (4,5).

Premise (4) seems to be analytically true, for no theory can explain what is not part of the phenomenon to be explained. Of course, one theory can explain another theory, but then that theory is not called a phenomenon. It may seem that McGinn has supported premise (5) with his argument that consciousness is not an observable feature of the brain. But if consciousness is an internal experience, a subjective phenomenon, then consciousness can be part of the data. Recall McGinn only showed that P was not an object of introspection, but consciousness itself is an object of introspection on his account. Hence McGinn's case for the epistemically closed nature of consciousness is problematic. If this criticism of McGinn is correct, then I have no good reason for refraining from drawing inferences about the nature of consciousness. McGinn's claim that consciousness cannot be explained in virtue of the spatial properties of the brain entails that it is not the case that both consciousness and the brain have spatial coordinates. I will argue that this concession entails that there is no physical explanation of the relationship between consciousness and the brain. If there is some explanation for this relationship, then there must be some nonphysical explanation of the relationship. Such a consequence means that consciousness is an emergent² property of the brain. The argument can be expressed in this way:²¹

7. If it is possible for there to be a physical explanation of a relationship between x and y, then both x and y have spatial coordinates.

8. The brain and consciousness do not both have spatial coordinates.

9. So it is not possible for there to be a physical explanation of a relationship between x and y.

McGinn concedes premise (8). Premise (7) is a conceptual truth. To be physical is to have spatial coordinates and vice versa. Empirical confirmation of the truth of premise (1) abounds. All explanations of relationships in physics or chemistry are explanations of processes or objects that have spatial coordinates. Any explanation of a relationship that purports to be physical without the thing to be explained having spatial coordinates stretches the boundaries of the physical beyond recognition.

McGinn might welcome this conclusion for he agrees that no explanation can be given of consciousness. But he cannot be so sanguine, for this conclusion precludes only physical explanations. Nonphysical explanations remain a possibility. McGinn has no argument precluding the possibility of such explanations. Furthermore, the argument above is due not to the limitations of our knowing power, but to the different kinds of properties characterizing physical reality and consciousness. The nonspatial character of consciousness conceded by McGinn is the ground for holding that there cannot be a physicalistic explanation of consciousness. That means consciousness cannot be simply a consequence of the interaction of the elements of the human organism. Since consciousness exerts causal power in giving meaning and explaining behavior McGinn's concession entails that consciousness is an emergent² property.

Of course this argument has no force for those who do not admit that consciousness cannot be explained by spatial processes in the brain.²² For them it is useful to call attention to the ability of human intelligence to know instantiable realities like validity, truth and equality. These instantiables do not occupy a position in space. If they did, they would no longer be instantiable.²³ Arthur Collins offers a similar argument based on the instantiability of beliefs. He distinguishes two notions of belief, (a) the temporally datable and essentially personal circumstance of someone's believing that p and (b) the propositional belief that p, in the sense of a belief that can be true or false and can be shared by many (or by none).²⁴ As a datable and localized event the first sense of belief could not be shared by others and hence is not the kind of entity that can be true or false. If we can know instantiable realities (and it seems that we do), then we have strong evidence for believing that human thinking is an emergent feature of the human being.

7. Conclusion and Suggestions

Although I have not proven that consciousness or thinking are emergent properties, I did respond to in-principle arguments against the very possibility of emergent properties. I have also shown that Searle and McGinn assign properties to consciousness or thinking that imply that thinking and consciousness are emergent properties.

But what could account for emergent features if it is not the interaction of the elements of which the whole is composed? The inability to answer this question, is what I think drives neurobiologists and philosophers to reject the reality of emergent phenomena.

Aquinas and Aristotle may be helpful here.²⁵ On their view what emerges is not just properties but a new substance, a new individual with a set of capacities that differ in kind from the capacities of the old substance or substances from which it comes to be.²⁶ Substances can be physical in the sense that they can be composed of physical parts. In this way the explanations provided by biochemistry and neurobiology are necessary to understanding human thinking. But as a whole the substance possesses a principle of organization that is not reducible to the interactions of discrete parts. As an organizing unity the substance has causal power.²⁷ On this view the emergent features are a function of the substance as an organizing unifying whole. For Aquinas and Aristotle this is true for plant and animal substances as well as for humans.²⁸ The necessity of higher-level biological explanations that do not reduce to lower-level microbiological explanations confirms the explanatory power of substances.

Substances are causal sources, but they are not themselves uncaused. I have argued that humans have an ability to relate to nonspatial realities. If this relationship requires a nonphysical explanation, then the nature of human substances would seem to require a nonphysical explanation as well.²⁹ The cause of human nature may turn out to be God for such a cause must be intelligent and lack spatial location.³⁰ However, much more would have to be done to argue that this agent is the God of Abraham. So philosophy of mind can give reasons for believing in a being that is intelligent and that lacks spatial location. Such a being might be the God of Judaism, Christianity or Islam, but it might simply be some other nonspatial intelligence. Philosophy of mind can contribute something to the

argument for God's existence, but it does not deliver the existence of the loving God revealed in Jewish, Christian or Islamic scriptures.

Endnotes

1. The belief in the literal seven-day creation of the world is inconsistent with data about the age of the earth and the fossil records of various species of plants and animals. For one excellent treatment of the difficulties with creation science see *Abusing Science: The Case Against Creationism* by Philip Kitcher (Cambridge, Mass: MIT Press, 1982).

2. This is certainly the view of J.J.C. Smart in "Sensations and Brain Processes," in *The Mind Brain Identity Theory*, ed. C.V. Borst (New York: MacMillan, 1970): 19-26. It is also the view of D.M. Armstrong in *A Materialist Theory of Mind*, (Routledge & Kegan Paul, 1968). Both of these authors claim as Searle has that since biology has already explained life in terms of physics and chemistry it is just a matter of time before consciousness and sensation are so explained. Colin McGinn is less explicit, but he holds a similar view. He believes that every phenomena that occurs in the universe is caused by the interaction of physical and chemical processes. He takes this view as a condition for the intelligibility of any natural phenomenon. See *The Problem of Consciousness* (Oxford: Basil Blackwell, 1991): p. 47.

3. Alfred North Whitehead, *Science and the Modern World* (New York: Macmillan, 1925) p. 52. The fallacy is the reification of the abstract determinism of physical theory onto the level of the concrete course of events. Talcott Parsons puts it this way, "Its starting point has been the possession of a general analytical scheme which, for a certain body of fact, works. This circumstance has been interpreted to mean, methodologically, that the concrete reality was reflected in the conceptual scheme, adequately for all scientific purposes. It has carried with it, inevitably, the implication of determinism. The logically closed system of theory becomes ... an empirically closed system," Parsons, *The Structure of Social Action*, p. 476.

4. Harry Austin, "Talcott Parson's Concept of Action" unpublished doctoral dissertation, 1975, p. 11. Austin takes his account from Arthur Pap's *An Introduction to the Philosophy of Science*, (Glencoe, Ill.: Free Press, 1962) p. 315.

5. Austin, p. 15.

6. It is surprising that McGinn confesses ignorance of the ancient and medieval arguments for God's existence based on the nature of thinking. He writes, "I do not know if anyone has ever tried to exploit consciousness to prove the existence of God, along the lines of the traditional Argument from Design, but in this post-Darwinian era it is an argument with more force than the usual one, through lack of an alternative theory. It is indeed difficult to see how consciousness could have arisen spontaneously from insentient matter; it seems to need an injection from outside the physical realm. Only something of the same kind could bring it about to begin with, it might be thought.

However, as will become clear, I do not really think we need to resort to God here ...,”
The Problem of Consciousness, p. 45.

7. John Searle, *Rediscovery of the Mind*, (Cambridge: Cambridge University Press, 1992) pp. 111-112. Searle makes the distinctions between causally emergent system features and radically emergent system features which I use in the text. In addition, he argues that it is possible that a property or feature of a system may be the effect of the interactions of the elements of that system in *Minds Brains and Science* (Cambridge, Mass.: Harvard University Press, 1984), pp. 21-22. He takes it as evident, for example, that liquidity, solidity and transparency are surface features caused by the behavior of elements at the micro-level as well as features of the very systems in question.

8. The basis for this argument is the following passage from Searle: “I cannot think of anything that is emergent², and it seems unlikely that we will be able to find any features that are emergent², because the existence of any such features would seem to violate even the weakest principle of the transitivity of causation,” *Rediscovery of the Mind*, p. 112.

9. Ernst Mayr, *Toward a New Philosophy of Biology* (Cambridge, Mass. Harvard University Press, 1988): pp. 8-21.

10. For a defense of this principle see Richard J. Connell, *Nature’s Causes* (New York: Peter Lang, 1995), pp. 132-134.

11. Thomas Nagel, “What it is like to be a bat?” in *Mortal Questions* (Cambridge: Cambridge University Press, 1979), pp. 165-180.

12. Nagel, p. 170.

13. Not all colors will be associated with a certain level of photon emissions. Some will be the result of the interaction of photon emissions.

14. Searle, *Rediscovery of the Mind*, pp. 122-123.

15. Searle, *Rediscovery of the Mind*, pp. 121-122.

16. A causal role of consciousness is again asserted in Searle’s argument that syntax is not intrinsic to physics. Searle argues that no computational state is discovered within the physics of a system. Anything intrinsic to physics is discovered within the physics of a system. Since any characterization of a process as computational is a characterization of a physical system from outside the system, the identification of a state as computational does not identify any intrinsic feature of the physics of the system. Rather, the characterization of a state as computational is an observer-relative characterization. According to Searle consciousness assigns the characterizations to systems. See Searle, pp. 209-212.

17. In Plato's dialogue, *Phaedo*, Simmias puts forth a view about the soul which is very similar to Searle's about consciousness (See *Phaedo* (92a-93b)). Simmias holds that the soul is a harmony of the bodily elements. The soul is not identified with any of the bodily elements but is identified with the harmony of the bodily elements. Socrates criticizes Simmias because his account of the soul renders the soul an effect and not a cause of bodily actions. Yet, Socrates avers, the soul does move the body, it fights off bodily temptations and so it cannot just be a function of the bodily elements. I criticize Searle in a similar vein. If his view of consciousness is as epiphenomenal as it seems given the trivial consequence that it is not reducible to a physical interaction, then how can he attribute to an appearance the power to generate meanings and interpretations of syntax as well.

18. McGinn notes that our thinking about mathematics and logic shows that we can form concepts in ways other than those based on introspection and perception. But these disciplines are not empirical ones. See p. 17 in McGinn.

19. See McGinn, p. 9 for another argument which can be expressed as follows:

1. If we can know the true psychophysical theory, then we are able to understand how the cat's brain generates consciousness. (pr)

2. If we are able to understand how the cat's brain generates consciousness, then it is possible for us to know the subjective form of the cat's experiences. (pr)

3. So if we can know the true psychophysical theory, then it is possible for us to know the subjective form of the cat's experiences. (1,2, HS) 4. It is not possible for us to know the subjective form of the cat's experience.(pr)

5. So we cannot know the true psychophysical theory. (3,4 MT)

It seems to me that premise (1) is true only if "true psychophysical theory" means the theory that explains any consciousness. But for a theory that just explained human consciousness and did not pretend to explain anything more, premise (1) is false.

It is not clear that the argument succeeds against a general theory of consciousness either, for premise (2) is problematic. That premise claims that understanding some naturalistic process implies something that is for us a logical impossibility, knowing the subjective form of the cat's experience. If premise (2) is read as an entailment or some kind of implication, then it is false, for it is possible that the antecedent be true and the consequent false.

20. McGinn, pp. 11-12.

21. This argument borrows a key premise from an argument about the consequences of being aware of instantiable universals presented by Thomas Sullivan and Russell Pannier in their forthcoming book, *Universals and the Immateriality of the Intellect*.

22. David Chalmers proposes that we take consciousness as an irreducible feature of reality and connect it to the basic physical realities like electrons and protons by a series of bridge laws. These laws would relate experience to elements of physical theory. See "The Puzzle of Conscious Experience" *Scientific American* (December, 1995): p. 83. But he assumes that consciousness is spatial for the analogy he makes with electromagnetic charge and gravitational fields are analogies with spatial realities. If consciousness is an irreducible nonspatial reality, however, his proposal will not succeed. On the other hand, if consciousness is spatial, his proposal has much to recommend.

23. This point about instantiatives can be found in Sullivan and Pannier's book, *Universals and the Immateriality of the Intellect*.

24. Arthur Collins, "Could Our Beliefs be Representations in our Brains?" *Journal of Philosophy* (1979) 76: pp. 225-43.

25. See Aristotle's *De Anima* in *The Complete Works of Aristotle* ed. by Jonathan Barnes (Princeton, New Jersey: Princeton University Press, 1984) and Aquinas' *Commentary on Aristotle's De Anima*, translated by K. Foster and S. Humphries (Notre Dame, Indiana: Dumb Ox Press, 1994).

26. A substance is not a modification of another entity like color is the modification of skin or surfaces. The existence of a substance is not dependent upon another as a property is dependent upon a subject. There are other kinds of dependencies that substance does exhibit. For example, substances do depend upon other conditions like oxygen, food and water. In the history of philosophy many have taken substance to be absolutely independent of everything. On this view only God is a substance.

27. In *The Human Person: Animal and Spirit* (Notre Dame, Indiana: University of Notre Dame Press, 1992) David Braine sets out an argument for the claim that neither humans nor animals are complex aggregates, but are in fact substances, unified wholes with intrinsic principles of action.

28. Eleanore Stump observes that in Searle's definition of emergent¹ features there is an important ambiguity in what counts as causal interactions of the parts. On the one hand such a phrase can simply refer to the interactions of the parts independently of their being parts of a whole. But on the other hand, the interactions may refer to the interactions the parts have in the form of the whole. In the latter sense the configuration or form of the whole is smuggled in. This last sense shows the causal role of substance as an organizing whole. If emergent¹ includes simply the first sense of causal relation, Searle may be mistaken to take liquidity as an emergent¹ property. Water may be an Aristotelian substance and not just an aggregate of oxygen and hydrogen. See "Non-Cartesian Substance Dualism and Materialism without Reductionism" by Eleanore Stump in *Faith and Philosophy* (October, 1995) 12: pp. 505-531.

29. The existence of the new nature requires this additional cause. The coming to be of the individual may be perfectly explained by biology.

30. Nothing in this argument requires abandoning the insights of evolutionary biology. The explanatory power of natural selection remains in tact. However, it is affirmed that these evolutionary explanations are not fully adequate explanations of human thinking. If God does cause human nature it is possible that he chose natural selection as the means whereby his own causality would be displayed. Natural selection could be an instrumental cause for the coming to be of human nature. Just as Michelangelo uses a chisel to be the instrument for his creation so too could natural selection have been God's instrument. See Connell, pp. 164-166 for more on accounting for new kinds of organisms.