

SEARLE'S AND PENROSE'S NON-COMPUTATIONAL FRAMEWORKS FOR NATURALIZING THE MIND

Napoleon M. Mabaquiao Jr.
De La Salle University, Manila

John Searle and Roger Penrose are two staunch critics of computationalism who nonetheless believe that with the right framework the mind can be naturalized. While they may be successful in showing the shortcomings of computationalism, I argue that their alternative non-computational frameworks equally fail to carry out the project to naturalize the mind. The main reason is their failure to resolve some fundamental incompatibilities between mind and science. Searle tries to resolve the incompatibility between the subjectivity of consciousness and the objectivity of science by means of conceptual clarification. He, however, fails to deal with the concepts crucial to this incompatibility, namely, the publicness of scientific knowledge and the privacy of psychological knowledge. Penrose tries to resolve the incompatibility between the non-computationality of psychological process and the computability of scientific process by expanding the scope of science through some radical changes in quantum physics. His strategy, however, has the danger of trivializing the distinction between science and non-science thereby putting into question the very value of the project to naturalize the mind. In addition, the feasibility of this strategy remains dubious in light of the mysteries that still surround quantum physics.

INTRODUCTION

The computational theory of mind (henceforth, computationalism) is one dominant framework for the naturalization of the mind or the assimilation of the mind into the scientific worldview. This framework is in fact what is adopted in *cognitive science*—the interdisciplinary scientific study of the mind. As Jay Freidenberg and Gordon Silverman (2006, 2-3) explain: “In order to really understand what cognitive science is all about we need to know what its theoretical perspective on the mind is. This perspective centers on the idea of computation, which may alternatively be called information processing” (see also Gardner 1985, 384-85; Hamish 2002, 2-7; Simon and Kaplan 1990, 2). Let us call the project to naturalize the mind the *naturalization project* and their proponents *naturalists*, while the project to carry out the naturalization project using the computational framework the *computationalist project* and their proponents *computationalists*.

It shall be observed that the failure of the naturalization project necessarily implies the failure of the computationalist project, but not vice versa; or the success of the computationalist project necessarily implies the success of the naturalization project, but not vice versa. In this consideration, we can divide critics of the computationalist project into two types: (1) those who believe that the mind is scientifically inexplicable and thus reject the feasibility of the naturalization project in all its possible forms, and (2) those who believe otherwise and thus maintain the feasibility of the naturalization project but only in its noncomputational form (that is, the use of a noncomputational framework to carry out the said project). We can call the former *nonnaturalists*, while the latter *noncomputationalists*. Nonnaturalists include the idealists, substance dualists, and natural mysterians;¹ whereas noncomputationalists include both nonrealist materialists,² under which we can classify the identity theorists, behaviorists, eliminative materialists, and instrumentalists, and realist materialists, under which we can classify the biological naturalists and proponents of Penrose's theory of mind—which I shall call the *quantum view of consciousness*.³ By our lights, computationalists and noncomputationalists are hence both naturalists—both subscribe to the naturalization project and they just differ as regards the appropriate framework to carry out the said project. Now while the success of the noncomputationalists in naturalizing the mind will necessarily prove the nonnaturalists wrong, their failure to do so will not necessarily prove the nonnaturalists correct (unless we grant their success in proving the computationalists to be wrong). Be that as it may, their (the noncomputationalists') failure to naturalize the mind will definitely strengthen the case of the nonnaturalists.

In this essay, I will examine the case of two noncomputationalists, namely, John Searle and Roger Penrose. What makes their case quite unique and interesting is that after arguing vigorously against the computationalist project both have advanced noncomputational frameworks to carry out the naturalization project. I argue that while both may be successful in showing why the computational framework will not work (in carrying out the naturalization project), they fail to show how their alternative frameworks in turn will. And this, in the main, is due to their failure to resolve or overcome some fundamental incompatibilities between science and mind. I will show that Searle fails to resolve an incompatibility arising from the nature of psychological and scientific knowledge, while Penrose fails to resolve an incompatibility arising from the nature of psychological and scientific processes. The essay is divided into two parts. In the first part, I put in proper perspective the views of Searle and Penrose by situating these views in the developmental stages of the naturalization project. In the second part, I examine the plausibility of their arguments for securing the possibility of the naturalization project.

The Naturalization Project and Computationalism

With computationalism as the reference point, the development of the naturalization project can be divided into the *precomputational*, *computational*, and *postcomputational* stages. These stages are distinguished primarily in terms of how the realization of the naturalization project is conceived. These stages are doctrinal and not historical in orientation, as some theories that will be classified under different stages may have been conceived in roughly the same historical period.

The Precomputational Stage. The precomputational stage is basically a reaction to *Cartesian dualism*, which divides reality into two qualitatively different types of substance: mind, the thinking but nonspatially extended substance; and matter, the spatially extended but unthinking substance. This dualism puts the mind outside the purview of science, thereby rendering a science of the mind impossible. For this reason, the precomputational theories of mind are bent on showing the mistake of Cartesian dualism, and on demonstrating that the mind, being a physical phenomenon, is very much within the compass of science. These theories argue for the nonexistence of the nonphysical Cartesian mind in two ways: by reducing mental phenomena to some form of physical phenomena, and by showing that the theory that postulates the existence of mental phenomena is either erroneous or held solely out of convenience or practical necessity.

Foremost of those that utilize the first method are the *identity theory*, which reduces mental states to neural states (see Smart 1991, 169-76), and *behaviorism*, which reduces mental states to behavioral dispositions.⁴ On the other hand, foremost of those that utilize the second method are *eliminative materialism* (of Paul and Patricia Churchland), which shows that the theory that postulates the existence of mental phenomena—called “folk psychology”—is wrong and outdated (see Churchland 1991, 601-12), and *instrumentalism* of Daniel Dennett (1991, 613-33), which shows that the attribution of mental states to an entity is just a convenient device for predicting its behavior. All these theories, after rejecting the existence of a nonphysical mind, redefine the concept of the mind in physical terms. While for behaviorists, the future science of the mind will be the same as a completed science of behavior, for the identity theorists, eliminative materialists, and instrumentalists, it will be the same as a completed science of the brain or neuroscience.

The Computational Stage. The computational stage develops as the computer technology is utilized in the pursuit of the naturalization project. This technology is not only presently the most sophisticated but proves to be powerful and flexible enough to simulate complex human cognitive processes (see Pylyshyn 1990, 52; Rumelhart 1990, 133). The result of this utilization is *computationalism* whose general thesis is that cognition is “a species of computing” (Pylyshyn 1990, 51) or, more specifically, that the mind is a kind of computer program that is realizable by appropriate pieces of computer hardware such as the human brain. Under this general thesis are the specific theses that human mental states and processes are computational states and processes, and that computers, believed to be capable of simulating human thought processes, are cognitive systems. In this stage, mental states are regarded neither as the states of some nonphysical substance nor as the physical states either of the brain or the external body, but as higher-level physical states realizable by the causal or functional organization of a physical system such as the computer and the human brain. In this regard, the computational stage is a reaction to both Cartesian dualism and the precomputational theories of mind.

Two disciplines are directly involved in the development of computationalism: philosophy and artificial intelligence. In the area of philosophy, the *functionalism* of Putnam (1991, 197-203), which basically grew out of the weaknesses of the identity theory and behaviorism, provided the impetus for the development of computationalism. The functionalist conception of the mind, however, was further solidified by the causal theory of mind developed by David Lewis (1991, 204-10) and D. Armstrong (1991, 181-88), according to which mental states are definable in terms of causal relations—that they are caused by some inputs and that they cause some outputs. There are two features of Hilary Putnam’s

functionalism that made the development of computationalism its natural consequence. The first is the *principle of multiple realizability*, which states that functional states are realizable in various physical systems that have the appropriate functional or causal organization. The second is the use of the concept of the *Turing machine*—the theoretical forerunner of the present-day digital computer—as the model for demonstrating the said principle (it is in this regard that Putnam’s functionalism is sometimes qualified as “machine functionalism”). Accordingly, as minds are like Turing machines, they can also be realized by inorganic or mechanical physical systems like the digital computers. This view culminated in the *language of thought hypothesis* of Jerry Fodor (1979), which argues that human cognition as a process of manipulating symbols uses a system of representation inherent in the human brain.

In the area of artificial intelligence,⁵ a subfield of computer science devoted to the construction of intelligent machines, the clearest expression of computationalism can be found in Herbert Simon and Allen Newell’s *physical symbol system hypothesis*, which regards intelligent systems as physical systems that manipulate symbols. Later on, two approaches to computationalism are distinguished: the *classical* (or the *symbolic*) *model* and the *connectionist* (or *artificial neural network*) *model*. The classical model, identified with Jerry Fodor, Zenon Pylyshyn, Herbert Simon, and Allen Newell, regards computing as symbol manipulation happening in a serial manner; while the connectionist model, identified with David Rumelhart, James McClelland, and Paul Smolensky, among others, regards computing as activations of (or the exchange of information among) the various units in neural networks happening in a parallel manner.

The Postcomputational Stage. As computationalism raises objections to precomputational theories of mind, postcomputational theories of mind in turn raise objections to computationalism. Postcomputational theories of mind, to begin with, share with computationalism the view that mental states are higher-level physical states; but they disagree with computationalism that these higher-level physical states are computational states. There are thus two sides to the arguments of the postcomputational theories: a negative side, where the weaknesses of computationalism are shown; and a positive side, where an alternative model for the naturalization project is advanced. In current literature, two postcomputational theories of mind stand out: Searle’s *biological naturalism* and Penrose’s *quantum view of consciousness*.

Searle’s and Penrose’s negative arguments hinge on a putative fundamental difference between the thinking process of humans and the computing process of machines/computers. Searle, through his *Chinese room argument* (see Searle 1980, 417-57), shows that this difference refers to the fact that human thinking process is inherently intentional, in that humans are aware of what their thoughts mean or represent in the world; while the computing process of computers is not, in that computers are not aware of what the symbols that they manipulate mean or represent in the world. Another way of saying this is that for human thinking both the semantics and syntax of its thoughts are necessary, while for the computing process of computers only the syntax of its symbols is necessary. As Searle (2004, 91) explains: “the computer operates by manipulating symbols. Its processes are defined purely syntactically, whereas the human mind has more than just uninterpreted symbols, it attaches meanings to the symbols.” Searle (2004, 92) later on also argues that the property of computability is observer-relative, meaning, computability is not an inherent property of things, even of computers, but an imposed one such that “you could not discover that the brain is a digital computer, because computation is not discovered in nature, it is assigned

to it” (see also Searle 1990). Consequently, it is trivial to say that the mind or the brain is a digital computer for anything (such as a wall or a pail of water) can be a digital computer if it can be described as implementing some computation or algorithm.⁶

On the other hand, Penrose, using insights derived from *Gödel’s incompleteness theorem*, shows that the putative difference refers to the fact that the human mind can transcend the rules of a formal system whereas the computer is necessarily bound by such rules. Penrose (1994, 64-65) writes: “. . . Gödel indisputably established was that no *formal* system of sound mathematical rules of proof can ever suffice, even in principle, to establish all the true propositions of ordinary arithmetic. . . his results showed something more than this, and established that human understanding and insight cannot be reduced to any set of computational rules.” To elaborate, a formal system, such as arithmetic, has propositions of two types. The first type refers to those whose truth is derivable from the rules of the formal system, and the second type refers to those whose truth is not derivable from the rules of the formal system. The human mind can recognize the truth of propositions of both types while the computer can only recognize the truth of propositions of the first type.

For their positive arguments, Searle turns to the discipline of biology while Penrose turns to that of physics. Accordingly, Searle’s *biological naturalism* argues that mental states are higher-level biological states whose properties (such as consciousness, qualia, and intentionality) are caused by the biological properties of the brain during the course of evolution. But though caused by these biological properties of the brain, mental properties, however, are not reducible to these same biological properties of the brain. Searle thus disputes the principle of multiple realizability, arguing that the biological makeup of the human brain is also essential for the production of mental properties. As Searle (2004, 113) writes: “Conscious states are realized in the brain as features of the brain system. . . .” In other words, it is important for the system that realizes conscious states to be a brain system. On the other hand, Penrose’s *quantum view of consciousness* argues that consciousness, together with other properties of the mind such as intentionality and qualia, arises from the *quantum activities* in the *cytoskeletal microtubules* in the neurons of the human brain. Penrose (1994, 367) remarks:

I am contending that the faculty of human understanding lies beyond any computational scheme whatever. If it is microtubules that control the activity of the brain, then there must be something within the action of microtubules that is different from mere computation. I have argued that such noncomputational action must be the result of some reasonably large-scale quantum-coherent phenomenon....

In this connection, a *revised quantum physics* is what is needed to scientifically explain the workings of the mind. It has to be quantum physics since mental states are quantum states of the brain, and it has to be a revised quantum physics to accommodate the noncomputational nature of mental states.

Securing the Possibility of the Naturalization Project

In telling us how they intend to carry out the naturalization project—as an alternative to the computationalist project—Searle and Penrose have not yet secured the possibility of this project. This is because they have yet to address the main obstacle to this project, namely,

that there is something fundamentally incompatible between mind and science. This incompatibility, on closer inspection, is precisely what has given rise to what has been called the “explanatory gap” by Joseph Levine (1983, 354-61) and the “hard problem” by David Chalmers (1995, 200-19) concerning the study of consciousness or the mind in general. This incompatibility comes in a specific form in the context of the respective frameworks proposed by Searle and Penrose. For Searle, it is how the subjectivity of consciousness can be studied using the objective methods of science. For Penrose, it is the noncomputational nature of how the mind works can be accommodated by science given the computational nature of its methods or procedures. On closer inspection, these two forms of incompatibility are closely related, if not interdefinable, for the subjective correlates with the noncomputational whereas the objective correlates with the computational. In what follows, let us look into how Searle and Penrose try to resolve the putative incompatibilities between mind and science.

OBJECTIVELY STUDYING THE SUBJECTIVE

After arguing that the computationalist project fails for leaving out the intentional feature of consciousness in its explanation of the workings of the mind, Searle proposes that consciousness be regarded as a higher-level biological phenomenon. This, however, does not yet address how consciousness given its subjective nature can be studied using the objective methods of biological science or of science in general. For his biological naturalism to be a viable alternative to the computationalist framework, he has to deal with this problem. Now Searle believes that he can resolve this difficulty simply by means of some conceptual clarification. Thus Searle (1999, 43) explains:

It is often argued that subjectivity prevents us from having a scientific account of consciousness, that subjectivity puts consciousness beyond the reach of scientific investigation. But typically, the argument rests on a bad syllogism. By exposing the fallacy in this syllogism, I believe we can come to understand subjectivity better. Here is how the argument goes:

1. Science is by definition objective (as opposed to subjective).
2. Consciousness is by definition subjective (as opposed to objective).
3. Therefore, there can be no science of consciousness.

Searle regards the above argument as a fallacy (particularly, an instance of *equivocation*) for containing ambiguous terms: the terms “objective” as ascribed to science and “subjective” as ascribed to consciousness. According to Searle’s analysis, these terms belong to different categories and are therefore *not direct opposites*. More specifically, the subjectivity of consciousness here, explains Searle, refers to the *kind of existence* attributed to consciousness; while the objectivity of science here refers to the *kind of knowledge* attributed to scientific knowledge. Since subjectivity refers to existence, Searle calls it “ontological subjectivity”; and since objectivity refers to knowledge, Searle calls it “epistemic objectivity.” Given these significations of the concepts *subjectivity* and *objectivity*, there is thus no contradiction in saying that there can be an objective study of a subjective phenomenon for what this really amounts to is that there can be *an epistemically objective study of an ontologically subjective phenomenon*.

In direct contrast to epistemic objectivity is, of course, epistemic subjectivity. As this dichotomy concerns knowledge, the question then is: What kind of knowledge is considered as subjective and what kind as objective? Searle (1999, 44-45) explains that if our knowledge is dependent on or is significantly affected by our attitudes and preferences, our knowledge is epistemically subjective; otherwise it is epistemically objective. A paradigm example of epistemically subjective knowledge is the kind of knowledge involved in evaluative statements. If I judge, for instance, that Baroque music is better than pop music, I do so because of my attitudes and preferences. In contrast, a paradigm example of epistemically objective knowledge is the kind of knowledge involved in descriptive or factual statements. If I say, for instance, that “*Jesu, Joy of Man’s Desiring*” was composed by Johann Sebastian Bach, I do so independent of my attitudes and preferences; that is to say, independent, for instance, of whether or not I prefer Baroque music to pop music. For whether I like it or not, such musical piece was composed by such composer. As scientific statements are factual and descriptive, such statements are thus epistemically objective.

On the other hand, in direct contrast to ontological subjectivity is ontological objectivity. And as this dichotomy concerns existence, the question then is: What type of existence is regarded as subjective and what type as objective? Searle (1999, 44-45) explains that the existence of something is subjective if it depends on some subject, while it is objective if it does not. The existence of conscious states is ontologically subjective in this regard since it is only meaningful to say that conscious states exist if there is some subject that has, experiences, or is conscious of them. For instance, pain and beliefs can only be said to exist if there is some subject that has or experiences them. It is absurd to say that there are pains and beliefs but no one has them. In contrast, the existence of physical and abstract entities is ontologically objective for it is meaningful to say that they exist even if there is no subject who is conscious of them. God, mountains, and chairs, for instance, can still be said to exist even if there is no subject who is conscious of them.

Based on these clarifications, it is thus clear why epistemic objectivity and ontological subjectivity are not direct opposites (and so are ontological objectivity and epistemic subjectivity). Each of these concepts belongs to a different category—the former to the category of knowledge while the latter to the category of existence. To understand them as direct opposites is thus to commit what Gilbert Ryle has called a *category mistake*. Given that scientific knowledge is epistemically objective while consciousness is ontologically subjective, Searle argues that there is nothing contradictory in having a scientific study of the nature of consciousness, for, again, what this really means is an epistemically objective study of an ontologically subjective phenomenon. Searle (1999, 45) explains:

So the fact that consciousness has a subjective mode of existence does not prevent us from having an objective science of consciousness. Science is indeed epistemically objective in the sense that scientists try to discover truths that are independent of anyone’s feelings, attitudes, or prejudices. Such epistemic objectivity does not, however, preclude ontological subjectivity as a domain of investigation.

The question, however, is whether this is really what the objectivity of science and the subjectivity of consciousness mean for those claiming that these two concepts are fundamentally incompatible. In the context of the significations attached by Searle to these

concepts, the incompatibility will arise only if these significations are attached to these concepts consistently. That is to say, in explaining away the incompatibility between these two concepts (the objectivity of science and the subjectivity of consciousness) by understanding one epistemically while the other ontologically, Searle supposes that those who believe that such incompatibility exists either understand both concepts epistemically or understand them both ontologically. More clearly, if Searle argues that there really is no incompatibility between A and B since A is actually X while B is actually Y, Searle supposes that the perceived incompatibility between A and B results from (mistakenly) regarding either A as X, and B as non-X, or A as Y, and B as non-Y. Now let us see whether Searle is correct in this supposition.

On the one hand, understanding both concepts epistemically (in Searle's sense) would mean that we understand a science of the mind as "an epistemically objective knowledge of an epistemically subjective phenomenon." The subjectivity of consciousness here would mean that our knowledge of consciousness would always be dependent on or would always be significantly affected by our attitudes and preferences; or that we can never have a factual or descriptive judgment about consciousness for our judgment about it would always be evaluative. This, however, does not seem to be what is at issue with regard to the subjectivity of consciousness. On the other hand, understanding both concepts ontologically would mean that we understand a science of the mind as "an ontologically objective knowledge of an ontologically subjective phenomenon." The objectivity of science here would mean that the existence of science or scientific knowledge is independent of some subject. Again, this does not seem to be what is at issue with regard to the objectivity of science. In light of these considerations, it is therefore dubious whether those who claim that there is a fundamental incompatibility between the concepts of objectivity of science and subjectivity of consciousness attach the same significations that Searle attaches to them, that is, epistemic and ontological significations.

There is, however, another type of significations that can be attached to the objectivity of science and subjectivity of consciousness not considered by Searle which gives rise to a fundamental incompatibility between these two concepts. This refers to the *public* nature of scientific knowledge and the *private* nature of psychological knowledge. (Generally, as this distinction also concerns knowledge this is also classified as an epistemological distinction; but this is different from the distinction made by Searle above between epistemic subjectivity and epistemic objectivity.) The private nature of psychological knowledge (knowledge of conscious states) refers to the fact that one can only have a direct knowledge of one's own conscious states.⁷ For instance, my knowledge of my own toothache is private since I am the only one who is directly knowledgeable about my own toothache; other persons' knowledge of my own toothache is merely indirect for it is based only on inferences from my verbal report and behavior. This is the sense in which consciousness is subjective. On the other hand, the public nature of scientific knowledge refers to the fact that the objects of this knowledge can in principle be directly known by everyone, or that in science what I know directly can in principle also be known directly by other people. For instance, if it is known in science that water is H₂O, this can be directly known by everyone. And this is the sense in which science is objective.

Another way of putting this sense of the objectivity-subjectivity distinction is as follows. The objectivity of science and the subjectivity of consciousness both concern *knowability*. The objectivity of science refers to the fact that the objects of

scientific knowledge are directly knowable by everyone; while the subjectivity of consciousness refers to the fact that conscious states are directly knowable only by the person who has them. Seen in this light, there is thus a clear inconsistency in saying that we can have an objective knowledge of something subjective. Saying that there is a science of the mind would mean here that we have a public knowledge about something we can only know subjectively. And needless to say, this is a contradiction.

Searle tries to dissolve the contradiction in having an objective study of a subjective phenomenon by showing that subjectivity and objectivity here belong to different categories. I have shown, however, that the categories that he considers, namely the epistemic and the ontological, are not really what are at issue. The concepts *subjectivity of consciousness* and *objectivity of science* is an issue but not because of the significations Searle attaches to these concepts. To make philosophical sense of this incompatibility, what is therefore needed is an understanding of the said concepts in light of another category. And this category, as I have shown, refers to the accessibility of knowledge, where the subjectivity of consciousness refers to the private nature of psychological knowledge and the objectivity of science refers to the public nature of scientific knowledge. Searle's conceptual distinctions fail to consider this category; as a result, his conceptual distinctions have failed to dissolve the incompatibility between the subjectivity of consciousness and the objectivity of science.

EXPANDING THE SCIENTIFIC DOMAIN

For his quantum view of consciousness to be a successful alternative to computationalism, Penrose still has to show how the noncomputationality of the mind can be explained using the computational methods of science. To fully appreciate the nature of this difficulty, we need to clarify that in saying that the method of science is computational, we mean that the scientific method proceeds according to step-by-step effective procedures.⁸ Given this, in saying that Penrose offers a noncomputational framework to naturalize the mind, we do not mean that the scientific method that he will use to explain the mind is noncomputational. What we mean, rather, is that his theory of mind regards the mind as noncomputational (that is, the mind does not proceed according to step-by-step effective procedures) but he nonetheless believes that we can have a scientific study of the mind. Thus the incompatibility arises: how can we account for something that does not proceed according to step-by-step procedures by a method that proceeds according to step-by-step procedures? Now Penrose thinks he can resolve this incompatibility by expanding our conception of science through some radical changes in quantum physics. What follows are three sets of remarks from Penrose to this effect:

[1] Does present-day physics allow for the possibility of an action that is in principle impossible to simulate on a computer? The answer is not completely clear to me, if we are asking for a mathematically rigorous statement. Rather less is known than one would like, in the way of precise mathematical theorems, on this issue. However, my own strong opinion is that such noncomputational action would have to be found in an area of physics that lies *outside* the presently known physical laws. (Penrose 1994, 15)

[2] For physics to be able to accommodate something that is as foreign to our current physical picture as is the phenomenon of consciousness, we must expect a profound change—one that alters the underpinnings of our philosophical viewpoint as to the nature of reality.... (Penrose 1994, 406)

[3] The conclusion is that whatever brain activity is responsible for consciousness (at least in its particular manifestation) it must depend upon a physics that lies beyond computational simulation. (Penrose 1994, 411)

In the first set of remarks, Penrose claims that what is needed is “an area of physics that lies *outside* the presently known physical laws.” In the second one, he says that physics has to undergo a “profound change—one that alters the underpinnings of our philosophical viewpoint as the nature of reality.” And in the third one, he says that this revised physics must be one which “lies beyond computational simulation.” He later clarifies that the revision that has to be made with quantum physics to account for consciousness will be the same revision that will be required of quantum physics in order to reconcile it (quantum physics) with the general theory of relativity. He (1999, xxii) writes:

I argue that a new theory will indeed be needed in order to make coherent sense of the “reality that underlies the stop-gap R-procedure that we use in present-day quantum mechanics [“?”], and I try to argue that it is in this undiscovered new theory that the required noncomputability will be found. I also argue that this missing theory is the same as the missing link between quantum theory and Einstein’s general relativity. The term used in conventional physics for this unified scheme is “quantum gravity.”

Penrose adds that he differs from most physicists who think that the required fundamental revisions to achieve quantum gravity have to be made only in the area of the general theory of relativity. For Penrose (1999, xxii), the fundamental revisions have to be made in the area of quantum mechanics as well. Surely, Penrose *cannot* agree with these other physicists, for if we grant the view of these other physicists then the changes that Penrose requires for quantum physics in order to accommodate the noncomputationality of mental states would most likely not be effected. For why should these changes be effected when they are not necessary to achieve quantum gravity?

We can identify at least two problems concerning the project of Penrose. The first concerns the consequence of the revisions required by Penrose for science to accommodate the noncomputationality of the mind. If science is to radically change such that what is at present considered nonscientific would later on become scientific, what happens in effect is that science extends its scope. It must be noted, however, that it is different when science extends its scope because of further scientific researches and when science extends its scope because it undergoes fundamental changes in its core principles. But in extending its scope because of radical changes in its principles surely it will not only be consciousness or mind that will be accommodated in its domain. In extending the scope of science to accommodate noncomputationality in its domain, the floodgates are, so to speak, opened. This would mean that other phenomena that do not presently fit into the scientific worldview, such as magic, paranormal phenomena or skills, and mystical experiences, in addition to

consciousness, would possibly be accommodated as well by the extended science. One critical consequence of this is the *demarcation problem*: how can science be so radically changed to accommodate the noncomputational and yet manage to retain its meaningful distinction from nonscience? As this problem threatens the general value of being scientific, it also questions the very point of the naturalization project. For what then would be the advantage of having a scientific understanding of the mind when science has weakened its standards, if not lost its rigor?

The second concerns the very nature of the revisions that Penrose requires of quantum mechanics to give room for the noncomputationality of the mind. Penrose, it will be recalled, claims that these revisions are the very same changes needed to achieve quantum gravity. But as Penrose himself pointed out, his idea that it should be in both areas of quantum mechanics and general theory of relativity that the scientific changes would have to be done parts ways with the idea of most physicists that such changes would have to be done only in the area of general theory of relativity. What this means is that Penrose still has to prove that the other physicists are mistaken in their hypothesis. The argument of Penrose is at best a *hypothetical* one. The radical changes required for quantum mechanics to pave the way for quantum gravity may indeed be our best *hope* for a science of the mind given that present-day science cannot account for the noncomputationality of the mind; but still this is just a hope, not a guarantee. We are still grappling with the mysteries of quantum mechanics—how to make sense of the world we live in given the findings in quantum mechanics. As Chalmers (1997, 333) writes:

The problem of quantum mechanics is almost as hard as the problem of consciousness. Quantum mechanics gives us a remarkably successful calculus for predicting the results of empirical observations, but it is extraordinarily difficult to make sense of the picture of the world that it delivers. How could our world be the way it has to be, in order for the predictions of quantum mechanics to succeed? There is nothing even approaching a consensus on the answer to this question.

That being the case, we do not yet know exactly how this theory of quantum gravity would be possible, much less how this theory would pave the way for a science of the noncomputational mind. As it is possible that this quantum gravity may not take place at all, it is equally possible that even granting that this quantum gravity is already in place, still we do not have a science of the mind.

CONCLUSION

The naturalization project has encountered various difficulties in each stage of its development. Either some type of incoherence arises, something essential about the mind is left out in the explanation, or the mysteries surrounding the mind remain. To date, there is no proposed science of the mind that has not encountered at least one of these forms of difficulties. What is perhaps needed are further ingenuity in theory building and further sophistication in our scientific tools. But all these will only matter if there is nothing fundamentally incompatible between science and mind. To secure the possibility of a future science of the mind, this incompatibility, first and foremost, has to be ruled out.

In this light, Searle and Penrose may be successful in demonstrating the weakness of the computationalism as a framework for the naturalization project, but their alternative noncomputational models can only be successful if they are able to overcome the putative fundamental incompatibility between science and mind. And we have shown that they are not able to do so. Searle tries to resolve the incompatibility between the subjectivity of consciousness and the objectivity of the scientific method by means of conceptual clarification. But he fails to consider the very antithetical concepts that have given rise to this incompatibility, namely, the publicness of scientific knowledge and the privacy of psychological knowledge. Penrose, on the other hand, tries to resolve the incompatibility between the noncomputationality of the mind and the computability of the scientific method by expanding the scope of science through some radical changes in quantum physics. Penrose's strategy, however, has the consequence of trivializing the distinction between science and nonscience, thereby undermining the very value of pursuing the naturalization project. Moreover, the feasibility of this project remains dubious in light of the mysteries that still surround quantum physics.

Finally, the question about the possibility of naturalizing the mind is not just a question of whether science will be able to complete its account of nature—as the mind is said to be the last piece in the grand puzzle. There is a larger question at stake. Our probing into the nature of the mind is precipitated, first and foremost, by our desire to understand who we are and to determine our proper place in the grand scheme of things. And we turn to science in the hope of giving rigor to the way we handle this inquiry. But given the failure of both computational and noncomputational models to naturalize the mind, perhaps it is not really the rigor of science that we need to have a deeper insight into the nature of our minds or of who we really are.

NOTES

1. For an explication of the position of natural mysterians, see McGinn (1997, 529-42). McGinn maintains that what will explain the nature of consciousness is some physical feature of the brain but he claims that such an explanation is not cognitively accessible to us. According to him, we are “cognitively closed” to such an explanation.

2. Generally, nonrealist materialists reject the nonphysical existence of mental phenomena and define the physical existence of such phenomena in terms of the neural states of the brain or the behavioral dispositions of the body. On the other hand, realist materialists also reject the nonphysical existence of mental phenomena but they also reject the view that the physical existence of mental phenomena is definable in terms of the neural states of the brain or the behavioral dispositions of the body. For according to the realist materialists, mental phenomena are higher-level physical phenomena.

3. Functionalists, who are definitely naturalists, can either be computationalists or noncomputationalists depending on the version being regarded. Computationalism is in fact regarded as just one form of functionalism, as it is sometimes also called “computational functionalism.”

4. This view is often associated with Gilbert Ryle (1965).

5. Among AI scientists who delved into the nature of the mind are John McCarthy, Marvin Minsky, Herbert Simon, Allen Newell, and Roger Schank.

6. For a good discussion of this point, see Jack Copeland 1996, 335-359.
7. This sense of the privacy of knowledge is what Ludwig Wittgenstein (1958) deals with in his famous *private language argument*.
8. Another way of saying this, based on the *Church-Turing thesis*, is that the scientific method is Turing-machine implementable.

REFERENCES

- Armstrong, D. M. 1991. The causal theory of mind. In *The nature of mind*. Edited by David M. Rosenthal. Oxford: Oxford University Press.
- Chalmers, David. 1997. *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
- _____. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2: 200-19.
- Churchland, Paul M. 1991. Eliminative materialism and the propositional attitudes. In *The nature of mind*. Edited by David M. Rosenthal. Oxford: Oxford University Press.
- Copeland, Jack. 1996. **That [not What?]** is computation? *Synthese* 108 (no issue no.?).
- Dennett, Daniel C. 1991. Three kinds of intentional psychology. In *The nature of mind*. Edited by David M. Rosenthal. Oxford: Oxford University Press.
- Fodor, Jerry. 1979. *The language of thought*. Cambridge: Harvard University Press.
- Freidenberg, Jay and Gordon Silverman. 2006. *Cognitive science: An introduction to the study of mind*. California: Sage Publications, Inc.
- Gardner, Howard. 1985. *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.
- Harnish, Robert M. 2002. *Minds, brains, computers: A historical introduction to the foundations of cognitive science*. Oxford: Blackwell Publishers.
- Levine, Joseph. 1983. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly* 64 (no issue no?).
- Lewis, David. 1991. Psychophysical and theoretical identifications. In *The nature of mind*. Edited by David M. Rosenthal. Oxford: Oxford University Press.
- McGinn, Colin. 1997. Can We Solve the Mind-body Problem? In *The nature of consciousness: Philosophical debates*. Edited by Ned Block et al. London: The MIT Press.
- Penrose, Roger. 1994. *Shadows of the mind: A search for the missing science of consciousness*. Oxford: Oxford University Press.
- _____. 1989. *The emperor's new mind: Concerning computers, minds, and the laws of physics*. Oxford: Oxford University Press.
- _____. 1999. New preface. *The emperor's new mind: Concerning computers, minds and the laws of physics*. Oxford: Oxford University Press.
- Putnam, Hilary. 1991. The nature of mental states. In *The nature of mind*. Edited by David M. Rosenthal. Oxford: Oxford University Press.
- Pylyshyn, Zenon. 1990. Computing in cognitive science. In *Foundations of cognitive science*. Edited by Michael Posner. Cambridge: The MIT Press.
- Rumelhart, David. 1990. The architecture of mind: A connectionist approach. In *Foundations of cognitive science*. Edited by Michael Posner. Cambridge: The MIT Press.
- Ryle, Gilbert. 1965. *The concept of mind*. New York: Barnes and Noble.

- Searle, John. 2004. *Mind*. Oxford: Oxford University Press.
- _____. 1999. *Mind, language and society: Doing philosophy in the real world*. London: Weidenfeld and Nicolson.
- _____. 1990. Is the brain a digital computer? Available at <http://cogsci.soton.ac.uk/~harnad/Papers/Py104/searle.comp.html>. Accessed: 2 May 2008.
- _____. 1980. Minds, brains, and programs. *Behavioral and brain sciences* 3 (no issue no.?).
- Simon, Herbert A. and Craig A. Kaplan. 1990. Foundations of cognitive science. In *Foundations of cognitive science*. Edited by Michael Posner. Cambridge: The MIT Press.
- Smart, J. J. C. 1991. Sensations and brain processes. In *The nature of mind*. Edited by David M. Rosenthal. Oxford: Oxford University Press.
- Wittgenstein, Ludwig. 1958. *Philosophical investigations*. Translated by G. E. M. Anscombe. Oxford: Basil Blackwell Ltd.

Submitted: 26 April 2013; revised: