The Journal of Mind and Behavior
Summer 1995, Volume 16, Number 3
Pages 235–254
ISSN 0271–0137

Unsolvable Problems, Visual Imagery and Explanatory Satisfaction

Marc F. Krellenstein

New School for Social Research

It has been suggested that certain problems may be unsolvable because of the mind's cognitive structure, but we may wonder what problems, and exactly why. The ultimate origin of the universe and the mind–body problem seem to be two such problems. As to why, Colin McGinn has argued that the mind–body problem is unsolvable because any theoretical concepts about the brain will be observation-based and unable to connect to unobservable subjective experience. McGinn's argument suggests a requirement of imagability — an observation basis — for physical causal explanation that cannot be met for either of these problems. Acausal descriptions may be possible but not the causal analyses that provide the greatest explanatory satisfaction, a psychological phenomenon that seems tied to the strength of the underlying observation basis but is affected by other factors as well.

Are there limits to the human mind or is its reach unlimited, able in principle to discover any and all facts of nature? The question has long been a subject of philosophical examination but has recently been posed as a psychological question: Are there problems that cannot be solved because of inherent limitations in human cognitive processing? Chomsky (1975), Fodor (1983) and Nagel (1986) have all argued for the existence of such unsolvable problems. Fodor says that these limitations are not only supported by his thesis for the modularity of various cognitive functions, but are almost certainly present in any case because of the existence at the lowest level of fixed and constrained cognitive structure:

I would like to thank Raymond Russ for his helpful comments on earlier drafts of this paper. This paper includes parts of a dissertation presented for the degree of Doctor of Philosophy at the Graduate Faculty of Political and Social Science of the New School for Social Research. The author is now affiliated with Ligature, Inc. Requests for reprints should be sent to Marc F. Krellenstein, Ph.D., 32 Duncklee Street, Newton, Massachusetts 02161 or by electronic mail to krellenstein@acm.org .

> Any psychology must attribute some endogenous structure to the mind . . . . And it's hard to see how, in the course of making such attributions of endogenous structure, the theory could fail to imply some constraints on the class of beliefs that the mind can entertain. (p. 125)

> A psychology which guarantees our epistemic unboundedness would thus have to guarantee that, whatever sort of subject domain the world turns out to be, somewhere in the space of hypotheses that we are capable of entertaining there is the hypothesis that specifies its structure . . . . I don't see how any remotely plausible cognitive theory could conceivably do so. (pp. 122–123)

Fodor offers support from an evolutionary perspective by observing that we accept such limitations without question in the case of other species, and "would presumably not be impressed by a priori arguments intended to prove (e.g.) that the true science must be accessible to spiders" (p. 126). The interesting claim in the case of humans is not that there is knowledge so beyond comprehension that we cannot even grasp the problems that such knowledge would address but rather that there are problems we can grasp but cannot solve. The psychological question of interest here is about problems where we suppose that the limitations on being able to solve them are conceptual, or cognitive, in nature. Such limitations would render the solutions literally incomprehensible to our minds, though not to some other possible mind, e.g., the mind of a more evolved species. This excludes from discussion problems whose limitations arise from formal aspects of the problem or surrounding theory, such as undecidable theorems in mathematics or quantum uncertainties, or from resource constraints such as our inability in a chess game to look ahead sufficiently far to determine the absolute best move.

The history of science and philosophy suggests at least two problems as being cognitively unsolvable: the problem of explaining the origin of the universe; and the "consciousness" part of the mind–body problem — providing an explanation for subjective experience.

The Origin of the Universe

Historical and modern-day attempts to explain the origin of the universe can be categorized into three groups: (1) single point of origin answers, (2) eternal universe answers, and (3) views that consider the problem unsolvable.

Single point of origin answers have in common the idea that the universe came into existence at some finite point in the past, before which there was "nothing," or at least no universe as we understand it. This sort of explanation can be seen in the Bible and other religious writings in which the universe came into being as a result of creation by a divine entity that is itself "uncaused" and eternal. A related line of thinking can be found in the modern-day classical Big Bang theory (see, for example, Hawking, 1988, or Penrose, 1989), which sees the universe as arising from a single momentous

explosion, a "singularity" outside the laws of science. Before this event there was simply "nothing," an emptiness consisting of neither space nor time.

Eternal universe answers postulate that the universe has existed forever, or perhaps is born and dies in an endless cycle. This was Aristotle's view in *On the Heavens* (McKeon, 1966), and can also be found in the modern-day "steady-state" theory, though this theory has few supporters today (see again Hawking, 1988). In such views every phenomenon or event admits of a prior causal explanation, extending indefinitely into the past and without any single origin point.

Views that consider the question of how the universe began as unsolvable see the question as scientifically unanswerable, "transcendent," or perhaps meaningless. Such views often include the position that the universe came into existence at a single point in time or perhaps existed forever, but consider such an origin or eternal universe as part of a larger problem to solve. The impossibility of either a finite or infinite universe was argued in Kant's first antinomy of pure reason (1787/1965). Contemporary philosophers such as Nagel (1986) have suggested that the things we cannot conceive "may include . . . what went on before the Big Bang" (p. 92). Munitz (1965, 1986) has been more emphatic, considering the problem "transcendent" and explicitly raising the possibility that there may be an answer that is beyond human comprehension. Wittgenstein (1921/1961) considered the question inherently unanswerable, but also raised the possibility of a "mystical" apprehension of the problem and a possible wordless solution, as shown in this well-known passage from the *Tractatus*:

> The solution of the riddle of life in space and time lies *outside* space and time . . . . The facts all contribute only to setting the problem, not to its solution. It is not *how* things are in the world that is mystical but *that* it exists . . . . Feeling the world as a limited whole — it is this that is mystical . . . . The solution of the problem of life is seen in the vanishing of the problem. (Is not this the reason why those who have found after a long period of doubt that the sense of life became clear to them have then been unable to say what constituted that sense?) There are, indeed, things that cannot be put into words. They *make themselves manifest.* They are what is mystical. [emphasis in original] (pp. 149–150)

As Wittgenstein observes, there is a central problem with trying to explain the universe: no matter what explanation any such answer provides, it seems we can still turn around and then ask for an explanation of *it* — what is *its* cause, or what came before *it*. In the case of a universe that has existed forever, this becomes the question of why or how there should be such a universe at all; and in the case of a universe arising from nothing, this becomes the question of why or how it was transformed from "nothing" to "something."

Simply put, we seem to be stuck with the problem of why there should be anything at all, whether that thing has always existed (eternal universe

answers) or came into existence at some point in time (single point of origin answers). The underlying problem is nicely captured by Gasking (cited in Black, 1964) in his commentary on Wittgenstein:

> What we demand as an answer is something like a well-confirmed hypothesis whose consequent is everything whatsoever — the world contemplated *sub specie aeterni* as a limited whole, limited by an antecedent which is something, in spite of everything being in the consequent. (p. 374)

Nozick (1981) makes the same point, noting that the difficulty makes the problem no less substantive:

> Any factor introduced to explain why there is something will itself be part of the something to be explained, so it (or anything utilizing it) could not explain all of the something — it could not explain why there is anything at all . . . . Some writers conclude from this that the question is ill-formed and meaningless. But why do they cheerfully reject the question rather than despairingly observe that it demarcates a limit of what we can hope to understand? (p. 115)

The force of the problem depends not only on our inability to identify an ultimate cause but also in our belief that there must be such a cause — our assumption that everything has a cause. Proponents of single point of origin or eternal universe answers might argue that this belief — or the related belief that an infinite series of causes must itself have an initiating cause — may simply be wrong, or at least admit of the origin of the universe as a single grand exception. The suggestion goes back at least as far as Hume (1739/1969). More recently, Smith (Craig and Smith, 1993) has observed that there seems to be no inherent logical contradiction about an uncaused entity (e.g., the Big Bang, or God) that is itself the initial cause of everything else.

While an uncaused origin seems technically possible, it seems unlikely that the idea that nothing happens without a cause, an idea which everywhere else is maintained, should be violated in the case of the origin of the universe (Craig makes a similar point in Craig and Smith, 1993). It seems more likely that our ability to discover and comprehend causality is, rather than irrelevant, inadequate for this problem. This does not rule out a resolution based on a wordless experience that, in Wittgenstein's terms, makes itself manifest. However, such a resolution would at a minimum be unlike any we have for any other scientific problem.

One alternative to an uncaused beginning is a solution based on the evolution of our concept of a cause of an entity. This might happen through discovery of as now unimaginable facts — Hawking (1988) raises the possibility that we could perhaps discover a theory so powerful that it compels its own existence (p. 174) — or through evolution of our ideas about causality or the universe.

Recently, just such a possible solution has been suggested based on work applying quantum mechanics to questions about the origin of the universe. This "quantum cosmology" provides a principled way to talk about "something coming from nothing," i.e., an explanation within the currently conceived laws of physics, and without resort to inexplicable "singularities." The idea is that a quantum fluctuation in the vacuum that preceded the universe led to the Big Bang and the subsequent creation of the universe, including the start of time itself. Grunbaum (1989) argues that such a cosmology has erased the question of the origin of the universe. This is because (a) the transition from the vacuum to the Big Bang is now explained by physical law, and (b) the period before the Big Bang is also before the start of space/time itself, and since there are no prior periods of time, there is no causation, making talk about a cause for the shift from the vacuum meaningless. Grunbaum states that Hawking "reaches the conclusion that there is no problem of creation, because at that stage, the very distinction between space and time becomes mushy . . . " (p. 393). Smith (Craig and Smith, 1993) has also argued that a quantum cosmology supports the idea of an uncaused beginning which needs no further explanation, and, in particular, no theistic explanation.

But this does not seem to square with Hawking's actual conclusions (1988). Despite his own endorsement of the "something from nothing" position Hawking states the following:

> How or why were the laws and the initial state of the universe chosen? (p. 173)
>
> Even if there is only one unified theory, it is just a set of rules and equations. What is it that breathes fire into the equations and makes a universe for them to describe? The usual approach of science of constructing a mathematical model cannot answer the questions of why there should be a universe for the model to describe. Why does the universe go to all the bother of existing? (p. 174)

This appears to suggest that the new quantum cosmology has, for Hawking, only added a layer of theory that itself needs to be explained. Although the universe can be seen to be caused by physical law and without a Creator, there is now the question of how such laws should exist. This seems to again point out that one can always demand a further explanation; that, as Munitz (1986) observes, the boundary of what is intelligible may shift but inevitably leaves an unintelligible residue; that, as Wittgenstein suggested, the facts only contribute to, and cannot solve, the problem of why there should be anything at all.

The Mind–Body Problem

How are we to explain the existence of subjective awareness, of one's sense of self, or of the "raw feels" or qualia that constitute sensations? Following Churchland (1984) we can quickly categorize the various historical and current approaches to solving this problem into three groups:

1. Dualist solutions that posit the existence of an independent subjective realm not reducible to physical or material phenomena.

2. Behavioral/linguistic solutions that effectively dismiss the problem as a pseudo-problem.

3. Materialist points of view that posit a physical/physiological substrate responsible for the existence of subjective phenomena. These include (a) identity theories that postulate a direct equivalence between brain states and mental processes; (b) forms of eliminative materialism that posit a neuroscientific basis for mental states that are, once understood, radically different from our common-sense understanding of those states as given by "folk psychology"; and (c) forms of functionalism, the computationally inspired view that equates mental processes or states (e.g., anger) with their place in a causal network that could potentially be instantiated by things other than human beings.

There is today relatively widespread agreement that subjective experience has some materialist basis, and one form or another of functionalism remains popular. But little progress has been made in providing any of the details of such an explanation. Wittgenstein (1953) expressed the difficulty as follows:

> The feeling of an unbridgeable gulf between consciousness and brain-process . . . . This idea of a difference in kind is accompanied by slight giddiness . . . . (p. 124)

Wittgenstein's explanation of the problem is linguistic confusion, a "logical sleight-of-hand" (p. 124) akin to the confusion shown by someone unfamiliar with universities who might visit each of the buildings in a university but still wonder exactly where the university itself was (the example is from Ryle, 1949). In this view minds are simply the wrong sort of thing to pose some of the problems we pose about them, and our mental states and desires are best viewed as dispositions and behaviors. But while such an analysis may help clarify certain aspects of the traditional mind–body problem, there seems to be a residual problem concerning the nature and origin by the brain of particular forms of subjective awareness. Linguistic analysis and the related psychological behaviorism that treated the problems of consciousness as fictions are not today generally seen to completely dissolve this problem (though some continue to argue otherwise; see, for example, Dennett, 1991).

But no other approach for bridging the gap between consciousness and brain-process has fared much better.

This is not to deny the progress that has been made in identifying physiological correlates of certain subjective experiences such as pain, or of the related lack of subjective experience in blindsight. However, such correlates do not seem to help in closing the "explanatory gap" between physiological process and subjective experience, in seeing just how the former actually gives rise to the latter. To close that gap appears to require bridging concepts that are fundamentally different in kind.

The counter responses to this apparent unsolvability that are contained within the other approaches to the mind–body problem can be viewed as similar to those offered to the position that the problem of the origin of the universe is unsolvable: we must either accept that consciousness is essentially uncaused or allow for the evolution of concepts and intuitions about the problem and/or new, currently unimaginable facts to somehow provide a solution. The first of these — viewing the problem as uncaused — is essentially what identity or various dualist positions do. Such theories reject the need for or existence of causal concepts that would allow physical brain phenomena to explain the nature of consciousness, arguing instead for entities that are equivalent to consciousness (identity theories) or have an independent existence of their own that is related to consciousness (dualist theories). This is not to say that consciousness is in these theories uncaused in just the way the universe may be considered to be uncaused. An identity theorist, for example, still thinks consciousness is embodied in and dependent on the brain. However, in both cases, there is a customary and expected level of causal understanding that is rejected as impossible or unnecessary.

The second counter response — that our concepts and intuitions may evolve or be changed by new, currently unimaginable facts — can be seen in the evolution of folk psychology envisioned by eliminative materialism. It can also be seen in the speculations by Nagel (1986) and Block (1993) that problem resolution is possible but would require concepts as yet unimaginable. It is difficult to argue against the possibility of discovering something currently unimaginable. The principal basis for such a possibility is the analogy with past scientific problems, but it is unclear just how good this analogy is.

Block claims to see a similarity between the difficulty of the mind–body problem and that of early attempts at understanding the physical basis of life, attempts which led to vitalism. It may be that the difficulty of the mind–body problem leads to an irreducible dualism or similar theory that, like vitalism, eventually gives way to a reductionist scientific explanation. But the earlier difficulty of finding explanations for aspects of life such as reproduction or purposeful behavior is arguably different from the conceptual

difficulty of finding brain processes that make intelligible the very character of subjective states. Churchland (1979) points out that when vitalism was popular "chemical theory already contained both the conceptual and the technical resources for a systematic attack on the problem of living tissue, construing it as a question of the chemical/structural/dynamical organization of matter" (pp. 109–110), and that the appeal of vitalism was less among those familiar with such chemical phenomena. By contrast, modern expertise in brain processes does not seem to provide either a model for seeing how subjective awareness could actually be explained by such processes or any general optimism about solving the problem.

A Possible Basis for Unsolvability

McGinn (1989, 1991) has presented a direct argument for the unsolvability of the mind–body problem. He suggests that any theoretical concept which serves to explain a property of the brain or other physical object must have its roots, if loosely, in perception. However, there can be no such concepts for explaining the brain's production of consciousness, since the property to be explained — consciousness — is itself paradigmatically unobservable. Any observation-based concept will therefore be unable to connect to consciousness in the way required of a full explanation of consciousness.

McGinn (1989) introduces the idea of observation-based concept formation as follows:

> Suppose we try out a relatively clear theory of how theoretical concepts are formed: we get them by a sort of analogical extension of what we observe. Thus, for example, we arrive at the concept of a molecule by taking our perceptual representations of macroscopic objects and conceiving of smaller scale objects of the same general kind. (pp. 358–359)

Such a theory of concept formation does not pertain to all abstract concepts but only to those concepts providing causal explanations of the properties of physical, material objects. Numbers, for example, do not seem to be such explanatory concepts. Numerical relationships can be seen to model and perhaps explain real-world phenomena in a process not fully understood, but we do not in any case accord them the kind of direct causation of physical phenomena we accord atoms or light waves. Nor is the model applicable to the explanation of non-physical phenomena such as the causes of World War II, or why a person chooses a certain hat to wear. Such phenomena reference social or intentional concepts as part of their explanation that are not necessarily derived from perception-based entities.

The problem, however, with trying to explain non-observable consciousness by reference to the brain phenomena we presume to be responsible for

it is that we are limited to observation-based concepts about the brain. No matter how much information about the brain we produce, the non-perceptual nature of conscious phenomena will preclude us from seeing how that information actually results in those phenomena. We might hope that some undiscovered concept will somehow overcome this limitation, but this is, for McGinn, little more than a belief in "magical emergentism," a willingness to believe in the possibility of new concepts that will magically escape the perceptual basis of all our observations and concepts about the brain. McGinn speculates that there may be other forms of intelligence for whom such concepts are possible, but only a belief in "magic" makes them plausible for humans.

We can elaborate on McGinn's suggestion of an observation basis for physical explanatory concepts and say that, if we assume this to be the only source of such concepts, it must then be possible to in some way visualize — to form an image of — any such concepts. The perceptual character of visual imagery is both a common-sense observation and one supported by a substantial empirical literature (see, for example, Kosslyn, 1980). By virtue of being imagable, such concepts are analogs of perceptual entities, even though the entities involved may not be directly observable. Thus, our understanding of liquids, to use an example from McGinn, is plausibly based on a molecular model that, though perhaps not observable, is itself based on a building-block model we can form an image of, and that can physically, if not observably, connect to the phenomena being explained.

There does not seem to have been any direct empirical investigation of McGinn's "clear theory" requiring an observation basis for concepts explaining physical phenomena. The idea does, however, have common-sense support, and may even be said to be a part of what we mean by physical causal explanation. For something physical to be a cause it is, after all, a thing in a way purely abstract concepts are not. Even "unobservable entities" are still entities. Larkin (1983) observed that the concepts used by experts in the representation of scientific problems, while more abstract than those of novices, nevertheless tended to have a perceptual, concrete basis:

> The naive representation [of the novice] is a direct simulation of events involving real (imagable) objects. It is less clear that the physical representation must always be imagable, but it is worthy of comment that most physical representations seem to have this feature. Even very abstract physical phenomena (e.g., energy states of an atom, conservation of quantum properties in the interaction of elementary particles) have corresponding imagable representations (energy levels, Feynman diagrams) used in solving related problems. (p. 79)

Larkin only says that "most" physical representations are imagable. However, her statement is not limited to causal concepts, so it does not, as far as it goes, contradict the claim that all causal concepts must be imagable.

McGinn's discussion includes not only the fact that the concepts are derived from observation but also that there must be some direct connection — some manner of spatial contiguity — between cause and effect. (Some form of temporal contiguity — Hume's "constant conjunction" — is presumably also a requirement, but not of immediate interest.) This will not be possible if the object to be explained is, like consciousness, non-observable. But even if the explanatory concepts and object to be explained are both observation-based it will do no good from a causal point of view if they are remote and related only by non-observation based laws. Absence of such locality is at least part of what lies behind the difficulties presented by what Einstein called the "spooky actions at a distance" of the Einstein–Podolsky–Rosen effect in quantum mechanics (Merwin, 1991, p. 502).

Applied to the problem of the origin of the universe, we can elaborate on the difficulty of establishing an initial cause by pointing to our inability to imagine — to literally form an image of — any constructs and process both apart from the universe and directly linking to it. The difficulty is that the object to be explained — the universe — contains all possible observable phenomena, forcing the explanation to be unobservable in principle. If our causal explanatory concepts are tied at least by analogy to what we can observe, we will be unable to form such concepts. There will be no observation-based concepts left to explain everything (the universe) that is potentially observable.

Such an elaboration does not negate our tracing the conceptual difficulty of explaining the universe to the very idea of causation as applied to the universe but extends and clarifies it by claiming that the mechanism of such problematic causal explanation is ultimately pictorial. The inability to explain the origin of the universe because of a lack of pictorial (imagable) constructs has in fact been suggested before. At the turn of the century physicist Ludwig Boltzmann urged that scientific thinking be limited to problems where mental pictures could be produced, stating that reasoning in the absence of such pictures was to "overshoot the mark" and was a waste of time, giving as an example the question of "why the world exists at all" (see Miller, 1984, p. 76). Such a requirement of imagability eliminates the possibility that a non-pictorial concept could be discovered that could be said to not be part of the universe but to still explain it; or, as Hawking suggested, that there could be a concept, presumably non-pictorial, so powerful it could cause its own existence.

So both the problem of the universe and the problem of consciousness can be viewed as presenting insurmountable obstacles to the attainment of a spatially based causal explanation. The former problem allows no spatial representation for an explanation that is not part of what is to be explained; and the latter precludes the existence of a common space between explanatory brain processes and subjective phenomena to be explained. Our own con-

sciousness is thus in some sense trapped within itself: it cannot look inward well enough to completely explain its own internal mechanics, and it is similarly limited in its outward reach to explain its ultimate origin.

The comparison also suggests a distinction between these two problems, though the conceptual difficulties of the problems makes any demarcation speculative at best. For the problem of consciousness we lack the concepts needed to close the "explanatory gap" between two domains of which we have some understanding: a spatially situated domain (brain processes) and a non-spatial one (conscious experience). Understanding the origin of the universe, however, requires us to escape an all-encompassing spatial domain that is everything we can or could causally understand. The need to "only" close an inter-domain gap for explaining conscious experiences suggests, as seems the case, that the problem may not be as widely perceived as unsolvable as the problem of the origin of the universe. Those who do consider the mind–body problem unsolvable may see the unsolvability as due to a mere mechanical limitation in brain concept formation. The difficulties of explaining the universe seem more fundamental, and may in fact pose explanatory difficulties beyond those under discussion.

McGinn's hypothesis of a required observation basis nevertheless unifies the problems by suggesting a sufficient and proximate source of the difficulty of each in the literal limitations of what we can imagine. Our experiences of these problems also support this similarity. Contemplation of both problems, though not experienced identically, is beset with conceptual confusion and "giddiness," and working scientists in both areas have generally tended to avoid focusing directly on them. New breakthroughs in cosmology or neuroscience are often announced together with renewed hopes of gaining on the problem of the origin of the universe or the problem of consciousness, respectively, but despite advancing the fields such breakthroughs seem to leave the respective problems untouched.

Explanatory Satisfaction

Flanagan (1992) has argued directly against McGinn's position, stating that McGinn is demanding too much of a potential scientific explanation of consciousness when he says that it must directly and completely reveal how the brain produces consciousness. Flanagan accepts that objective physical explanation cannot capture subjective conscious phenomena and states that there are good reasons for this because of the way consciousness is "hooked up." However, he argues, that does not preclude us from finding a complete physical explanation as good as other accepted scientific explanations, e.g., explaining the chemical properties of water from its chemical structure as H_2O. Flanagan asserts that none of these accepted explanations is ever

completely satisfying. Hardcastle (1993) goes further, stating that first-person accounts and third-person accounts are potentially "just different sorts of descriptions of the same events." She claims that nothing is lost from a scientific explanation of consciousness "if both conceptual frameworks can describe the same causal interactions, albeit in different terms" (p. 32).

But it is hard to see in what way H_2O is significantly inadequate as a scientific explanation of water. As far as H_2O explaining the other properties of water, these can be derived from the properties of water's constituent elements with a directness which Flanagan and Hardcastle concede that physical explanation could not provide for subjective consciousness. It is true that water being H_2O offers no clue about water's phenomenally experienced qualities, or about why water is H_2O and not something else (Flanagan refers to the "contingent" nature of H_2O as an explanation). However, the former is the consciousness problem, and the latter is the problem of why things are what they are and exist as they do at all — arguably, the origin of the universe problem. These problems indeed continue to exist within scientific explanations. However, one cannot use the failure to solve them as evidence for a diminished standard of explanatory adequacy for these very problems when other aspects of scientific phenomena are well explained.

Even if Flanagan and Hardcastle were to accept that potential explanations of consciousness are not only limited but must fall short of other scientific explanations they might still say they could be "fully satisfied" with such explanations precisely because those are the best or only explanations that are possible. It might, indeed, be argued that Dennett's *Consciousness Explained* (1991) establishes him as one already satisfied that the problem has been solved. The situation for accepting such physical explanations of consciousness is similar to the possibility of accepting the idea of an uncaused universe. In both cases, we seem to have (a) unsolvable problems or, at best, explanations unlike any other satisfactory explanations, and yet (b) at least some individuals (e.g., Flanagan and Grunbaum, respectively) who claim to be satisfied by such actual or potential explanations.

Even if we accept that the problems of consciousness and of the origin of the universe are unsolvable or admit only of explanations unlike other scientific explanations, claims of explanatory satisfaction are nevertheless possible because the satisfaction an explanation provides is at least partly a psychological question — a judgment by the individual of the adequacy of some internal state of comprehension. As a psychological judgment we would expect satisfaction to be determined not solely by the characteristics of the problem and proposed solution but by cognitive structure and relevant individual differences. We can try to sketch how such psychological factors might be involved.

We should note first that there are probably various levels of explanatory satisfaction and not simply a binary choice between perfect and imperfect

intelligibility. Starting with interactions between real, tangible objects, it seems that many directly observable physical events, such as the movement of billiard balls after a collision, are immediately apprehended and causally understood, at least at some macro level of analysis. Such mechanistic phenomena seem to appeal to a quickly developed or possibly hard-wired perception and a grasp of physical causation that is highly satisfactory. As part of our "naive physics," such recognition makes a verbalized or abstract (and philosophically controversial) explanation unnecessary. Leslie and Keeble (1987) found that sensitivity to this sort of directly observed causal connection occurs in infants as young as six months. The Gestalt "common cause" illusion would also seem to be evidence for some innate ability to recognize causation. The origins of such forms of recognition and understanding are presumably found in the direct evolutionary benefits these forms confer for survival in the world.

The existence of a molecule or electron and its place in various explanations does not have the simplicity, immediacy or concreteness of a mechanical collision among billiard balls, and to that extent is likely to count as less satisfactory. The abstractness of such objects and explanations comes from a direct resource limitation — they are simply too small to view. We can, however, still imagine, from analogy with observed objects, what such objects must be like and how they might interact.

Abstractions involving waves, fields and other similar entities seem further removed from these imagined mechanical interactions between very small but discrete objects, and explanations involving these abstractions are less satisfactory as explanations. These abstract objects nevertheless remain imagable to some degree, even if the analogy with observables (ocean waves, concentric ripples in a pond, etc.) becomes more tenuous. There are also abstract attributes or "causal powers," such as spin or attraction, that may apply to such entities. These concepts may not themselves be imagable but apply to entities that are, and would not seem to have any causal role independent of those entities.

With such complex abstractions we might also expect to see an increase in variation of actually experienced levels of satisfaction owing to individual differences. One might reasonably assume, for example, that physicists or others practiced in such matters would come to feel more comfortable with these entities, if not as comfortable as with directly perceived objects. DiSessa (1983) in fact observed that one difference between experts and novices is precisely their repertoire of and familiarity with such abstractions. For experts, these abstractions tend to become learned phenomenal primitives, recognized and understood with at least some of the immediacy accorded everyday objects.

We might place other abstractions at other points on a continuum of intelligibility or level of subjective satisfaction, but the general idea is that we have a range of levels of satisfaction determined by how removed the abstractions are from direct and hard-wired recognition. Some modification of those levels is probably possible as a result of expertise or habituation. But, as Larkin observed, the objects involved in most if not all of these and many other possible examples of good scientific explanation are, if not directly observable, analogs of observable entities. The more abstract and less satisfying objects are perceptually more remote but still imagable.

McGinn's suggested requirement of an observation basis for explanations of physical phenomena can now be seen as a hypothesis for the basis of such explanations. The low-level recognition of causation, though enriched or overruled by cognitive processes, serves as the template for causal analysis. The degree of satisfaction of a given explanation is then determined by the strength of that observation basis — the fit of the explanation to the form of the underlying perceptual template. This is itself a factor both of some measure of psychological distance from the original concrete objects and possibly other individual differences, such as expertise, that can affect the perceived phenomenal character of the explanation as more or less closely tied to the concrete world.

Tying explanatory satisfaction to the perception of concrete objects or to those derived from them would seem to fly in the face of the historically increasing abstractness and mathematicization of science. Kuhn (1977) for one has argued that what counts as explanatory in science has evolved over the centuries. He cites a movement from innate properties to mechanical interactions, and then to the mathematical, probabilistic and even indeterminate forms of contemporary physics, which, according to Kuhn, usually eschews references to causes altogether. This suggests that pinning causal explanation to observation-based interactions is a reactionary effort that ignores progress in the very structure of causal explanation.

But the acausal explanations Kuhn claims have superseded these observation-based causes would seem to be just that — acausal explanations/descriptions but not causal analyses/explanations. Cartwright (1983) has argued that working physicists have not given up their need for a single causal story for a given phenomenon, and treat what are often redundant applicable laws as merely practical means of computation. Cartwright admits the value of such instrumental laws (e.g., f=ma) as explanations. But far from allowing an empiricist reconstruction of causation in terms of such explanatory laws (see Gasper, 1991, for an overview on such reconstruction efforts), she complains that such instrumental explanations have been confused with real causal explanation since the time of Aristotle. She argues further that it is only causal explanation that offers a true or false versus a merely instrumental account of a phenomenon.

An observation basis for physical causal explanation does not explain everything about what makes an explanation causal, which is a long-standing philosophical problem, but states a requirement for the psychological perception of causation. The strength of the observation basis determines the satisfaction of the causal explanation, and where such a basis does not exist at all, only acausal explanations will be possible. Such acausal explanations will typically not make reference to observation-based entities but will explain by reference to mathematical laws or other relationships. Acausal explanations might also involve observable entities, such as brain states, but will connect them to what is being explained, e.g., consciousness, only by association/correlation and without the contiguity or directness of causal explanation.

As Cartwright observes, such acausal explanations are properly seen as distinct from causal explanations. When the two forms are conflated we would expect that the above criterion of satisfaction would render acausal explanations the least satisfying of all. When the two forms are seen as distinct, and as meeting different goals and standards, one would think an acausal explanation could be completely satisfying as far as it goes. Still, one might expect that the more limited goals of acausal explanation and the lack of a possibly hard-wired perceptual substrate would, all things being equal, amount to a less satisfying experience of comprehension and problem resolution.

However, the contrasting views on the problems under discussion suggest (if suggestion were needed) that all things are not equal. In the case of the problem of consciousness and the views of McGinn and Flanagan, it is hard to distinguish among differences in satisfaction with acausal explanations, optimism over finding a causal explanation and disagreement about the constitution and satisfaction of causal explanations — though probably all play a part. In particular, some people may rate acausal explanation as highly satisfactory precisely because they believe it is the best they can get in a given situation, and ignore, discount or are even desensitized to the greater psychological satisfaction of a causal explanation.[1] For the problem of the origin of the universe the contrast is often more clearly over acceptance versus rejection of an acausal explanation as the only explanation available to us. This is perhaps because the possibility of a full causal explanation seems more remote for this problem than for the problem of closing the inter-domain gap between brain processes and consciousness. Such causal explanations for the

[1]Another relevant individual difference may be one's use of visual imagery. Less frequent (or less vivid) imagers may be less inclined to perform visual enactments of causal processes, making them less sensitive to the difference between causal and acausal explanations and more accustomed to experiencing all explanatory concepts as arbitrarily abstract. Such individuals may therefore be more satisfied with acausal explanations, or more optimistic about finding unspecified future explanations. Further discussion of this idea and some preliminary support for it from interviews with physicists can be found in Krellenstein (1994).

universe as are embraced or hoped for are often religious in nature, though a religious explanation only seems to push the causal problem back a step. Smith, who accepts the idea of an uncaused universe, concludes his book-long debate (Craig and Smith, 1993) with Craig's theistic account of the origin of the universe by highlighting this difference over acceptance or rejection of an acausal explanation:

> There is an underlying agreement in attitude that motivates Craig's and my various efforts to fathom the universe's existence, namely, a wonder or awe that there is not nothingness . . . . Craig adds that this "astonishment should not end in a mute stupefaction but lead us . . . to the intelligible explanation of the universe." The considerations adduced . . . suggest that we may agree on this point as well, with the difference between us coming down to the question: Is the intelligible explanation of the universe causal or acausal? (p. 337)

Conclusions

There are limits on human knowledge. This seems at least as certain as our belief that all other forms of life have even more limited understanding of the world than we do (or none at all). More specifically, we seem to have good reasons for thinking we cannot solve the mind–body problem or the problem of the origin of the universe. If this is so it would not preclude us from finding acausal relationships that describe one or another aspect of these problems, but it would deny us the causal explanations we strive for in understanding the world and that, all things being equal, provide our deepest sense of comprehension and explanatory satisfaction. The unsolvability may be rooted in the perceptual basis of our casual understanding, which evolved for the workaday tasks that have shaped our survival but is not quite up to the problem-raising ability of our minds. This is depressing, but should not be surprising. The evolutionary benefits of our cognitive capacities and curiosity must over time enhance our reproductive potential, but they do not require that all conceivable problems be solvable any more than they should lead automatically to individual happiness (see Wright, 1994, for a recent discussion of the mismatch between what evolution provides and what we may choose to consider important).

We cannot rule out the possibility that consciousness and the universe are simply uncaused — that there is no more to the story than all the acausal descriptions we have or could discover — but it seems unlikely based on what we do know about the world. We can also not rule out the possibility of discovering concepts as yet unimaginable that well help us solve these problems, but the possibility has little more than faith to recommend it. Should we encounter a higher intelligence (or eventually evolve into one ourselves) we might be able to put both these propositions to the test. The likely outcome, it has been argued, is that we would find out that such an intelligence

has more to say about these problems than we could ever discover on our own, but we would not be able to understand what it would say about them. (Finding out there is something more to say about these problems is not only likely but probably desirable, since causal explanations that we cannot discover or understand may be easier for most people to grasp and accept than no causes at all.)

Accepting such substantive limits on our understanding induces a fitting humility and continues a realistic re-adjustment of our view of ourselves that began at least as long ago as Copernicus. Scientists working in these areas will need to be content with such acausal descriptions as can be developed. This should temper the renewed optimism and talk about solving these problems that seem to accompany every significant advance in their respective fields. It should also curtail the radical, often desperate attempts to fashion a causal story where none is to be had, e.g., Penrose's (1989) attempt to explain consciousness through quantum mechanics. The strangeness of some of these theories is understandable given the difficulty of the problems. With regard to the origin of the universe problem, Nozick (1981) says:

> The question cuts so deep . . . that any approach that stands a chance of yielding an answer will look extremely weird. Someone who proposes a non-strange answer shows he didn't understand the question. Since the question is not to be rejected, though, we must be prepared to accept strangeness or apparent craziness in a theory that answers it. (p. 116)

Nozick himself then goes on to suggest several strange candidate explanations, none of which appears to have moved the problem any closer to solution.

Accepting the unsolvability of the mind–body problem also has implications for the much-debated if ill-defined question of whether a machine could be conscious, at least in the sense of possessing the subjective awareness that has been the aspect of consciousness under discussion. These implications require more discussion than is possible here but the basic idea is that our lack of understanding of how consciousness could arise from the brain or any other material precludes any easy answer to the question of machine consciousness or to related questions of the role of organic materials, transducers, analog processes, etc. in producing such consciousness. We cannot see how any amount of computer syntax could result in subjective understanding or awareness (e.g., Searle, 1984) but we also cannot see how any amount of biological process could produce these phenomena. In the latter case we have the brute fact that it is somehow done, but without knowing how we cannot say what the crucial mechanisms are and how they might be instantiated. We may know more about computers "from the ground up" than we do about the relevant biology but it is unclear if that should lead us to be pessimistic about a computational (but still ultimately inaccessible)

theory of consciousness given that any and all physical theories seem blocked off from explaining consciousness. While the unsolvability of the mind–body problem may not eliminate the possibility of machine consciousness it does prevent us from ever knowing exactly how to produce it, except perhaps as an indirect byproduct of some more global construction that produced consciousness by mechanisms that are themselves unknown.

For some, this latest surrender on what can be known about consciousness and the universe will be too much, and resisted with optimistic exhortations not to give up as much as with more direct counter-arguments. Flanagan has labeled the position of McGinn and others as the "new mysterianism" (1992, p. 109; "new" mysterianism to indicate the position accepts a naturalist explanation but declares it unattainable, versus older, anti-naturalist views on the impossibility of explanation), and is openly troubled that declaring the mind–body problem unsolvable will lead us "into not trying to understand mind" (p. 128). This seems unlikely, since there is much of interest to be learned that can be learned, but such a declaration may certainly influence some individuals (mostly philosophers?) to pursue more tractable problems than the ultimate connection between conscious experience and brain process or the ultimate origin of the universe. Whether this is good or bad depends mainly on one's view of the merits of the arguments for unsolvability, though it can be granted that there are risks in abandoning investigations while the arguments are less than conclusive.

To the extent the meaning of our life is based on understanding its ultimate origin or the nature of our own sense of self, the unsolvability of these two problems further suggests that our lives may never be completely based on or derive their meaning from facts that have the particular certainty that scientific knowledge has for us. For understanding both ourselves and the universe we will need to try to be content with less decisive considerations, futile as that effort may be for some of us.

References

Black, M. (1964). *A companion to Wittgenstein's "Tractatus."* Ithaca: Cornell University Press.

Block, N. (1993). Review of *Consciousness Explained. Journal of Philosophy, 90,* 181–193.

Cartwright, N. (1983). *How the laws of physics lie.* New York: Oxford University Press.

Chomsky, N. (1975). *Reflections on language.* New York: Pantheon Books.

Churchland, P. (1979). *Scientific realism and the plasticity of mind.* Cambridge, England: Cambridge University Press.

Churchland, P. (1984). *Matter and consciousness.* Cambridge, Massachusetts: MIT Press.

Craig, W.L., and Smith, Q. (1993). *Theism, atheism, and big bang cosmology.* New York: Oxford University Press.

diSessa, A. (1983). Phenomenology and the evolution of intuition. In D. Gentner and A. Stevens (Eds.), *Mental models* (pp. 15–33). Hillsdale, New Jersey: Erlbaum.

Dennett, D. (1991). *Consciousness explained.* Boston: Little, Brown.

Flanagan, O. (1992). *Consciousness reconsidered.* Cambridge, Massachusetts: MIT Press.

Fodor, J. (1983). *The modularity of mind.* Cambridge, Massachusetts: MIT Press.

Gasper, P. (1991). Causation and explanation. In R. Boyd, P. Gasper, and J.D. Trout (Eds.), *The philosophy of science* (pp. 289–297). Cambridge, Massachusetts: MIT Press.
Grunbaum, A. (1989). The pseudo-problem of creation in physical cosmology. *Philosophy of Science, 56*, 373–394.
Hardcastle, V.G. (1993). The naturalists versus the skeptics: The debate over a scientific understanding of consciousness. *The Journal of Mind and Behavior, 14*, 27–50.
Hawking, S. (1988). *A brief history of time*. New York: Bantam Books.
Hume, D. (1969). *A treatise of human nature*. London: Penguin Books. (Original work published 1739)
Kant, I. (1965). *Critique of pure reason* [N.K. Smith, Trans.]. New York: St. Martin's Press. (Original work published 1787)
Kosslyn, S.M. (1980). *Image and mind*. Cambridge, Massachusetts: Harvard University Press.
Krellenstein, M.F. (1994). Visual imagery and the limits of comprehension (Doctoral dissertation, New School for Social Research). *Dissertation Abstracts International, 55*, 1172B.
Kuhn, T.S. (1977). Concepts of cause in the development of physics. In T.S. Kuhn (Ed.), *The essential tension* (pp. 21–30). Chicago: University of Chicago Press.
Larkin, J.H. (1983). The role of problem representation in physics. In D. Gentner and A. Stevens (Eds.), *Mental models* (pp. 75–98). Hillsdale, New Jersey: Erlbaum.
Leslie, A.M., and Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*, 265–288.
McGinn, C. (1989). Can we solve the mind–body problem? *Mind, 391*, 349–366.
McGinn, C. (1991). *The problem of consciousness*. Oxford: Blackwell.
McKeon, R. (1966). *The basic works of Aristotle*. New York: Random House.
Merwin, N.D. (1991). Is the moon there when nobody looks? Reality and the quantum theory. In R. Boyd, P. Gasper, and J.D. Trout (Eds.), *The philosophy of science* (pp. 501–516). Cambridge, Massachusetts: MIT Press.
Miller, A.I. (1984). *Imagery in scientific thought*. Cambridge, Massachusetts: MIT Press.
Munitz, M.K. (1965). *The mystery of existence: An essay in philosophical cosmology*. New York: Appleton–Century–Crofts.
Munitz, M.K. (1986). *Cosmic understanding*. Princeton: Princeton University Press.
Nagel, T. (1986). *The view from nowhere*. New York: Oxford University Press.
Nozick, R. (1981). *Philosophical explanations*. Cambridge, Massachusetts: Harvard University Press.
Penrose, R. (1989). *The emperor's new mind*. New York: Oxford University Press.
Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
Searle, J. (1984). *Minds, brains and science*. Cambridge, Massachusetts: Harvard University Press.
Wittgenstein, L. (1953). *Philosophical investigations* [G.E.M. Anscombe, Trans.]. New York: Macmillan.
Wittgenstein, L. (1961). *Tractatus logico-philosophicus* [D.F. Pears and B.F. McGuinness, Trans.]. New York: Routledge and Kegan Paul. (Original work published 1921)
Wright, R. (1994). *The moral animal: The new science of evolutionary psychology*. New York: Pantheon Books.