

Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False

Thomas Nagel

Print publication date: 2012

Print ISBN-13: 9780199919758

Published to Oxford Scholarship Online: Jan-13

DOI: 10.1093/acprof:oso/9780199919758.001.0001

Cognition

Thomas Nagel

DOI: 10.1093/acprof:oso/9780199919758.003.0004

Abstract and Keywords

This chapter examines another type of antireductionist argument and its consequences. It first identifies the problem concerning mental functions such as reasoning, thought, and evaluation that are limited to humans. It questions whether a person's cognitive capacities can be placed in the framework of an evolutionary theory that is no longer exclusive but still retains the Darwinian structure. It then shows ways of thinking about the basic reasoning capacities and studies the faculty that allows people to escape from the world of appearance that is presented by their prereflective innate dispositions, and move into the world of objective reality. This chapter also takes a look at teleological principles as part of the natural order and lists some alternatives.

Keywords: antireductionist argument, mental functions, cognitive capacities, evolutionary theory, Darwinian structure, basic reasoning, world of appearance, objective reality, teleological principles, natural order

1

I now want to take up a different type of antireductionist argument and its consequences. Consciousness presents a problem for evolutionary reductionism because of its irreducibly subjective character. This is true even of the most primitive forms of sensory consciousness, such as those presumably found in all animals. The problem that I want to take up now concerns mental functions such as thought, reasoning, and evaluation

that are limited to humans, though their beginnings may be found in a few other species. These are the functions that have enabled us to transcend the perspective of the immediate life-world given to us by our senses and instincts, and to explore the larger objective reality of nature and value.

I shall assume that the attribution of knowledge to a computer is a metaphor, and that the higher-level cognitive capacities can be possessed only by a being that also has consciousness (setting aside the question whether their exercise can sometimes be unconscious). That already implies that those capacities cannot be understood through physical science alone, and that their existence cannot be explained by a version of evolutionary theory that is physically reductive. But the problem I now want to discuss goes beyond this. It has to do with the nature of these capacities and the relation they put **(p.72)** us in to the world. What we take ourselves to be doing when we think about what is the case or how we should act is something that cannot be reconciled with a reductive naturalism, for reasons distinct from those that entail the irreducibility of consciousness. It is not merely the subjectivity of thought but its capacity to transcend subjectivity and to discover what is objectively the case that presents a problem.

Thought and reasoning are correct or incorrect in virtue of something independent of the thinker's beliefs, and even independent of the community of thinkers to which he belongs. We take ourselves to have the capacity to form true beliefs about the world around us, about the timeless domains of logic and mathematics, and about the right thing to do. We don't take these capacities to be infallible, but we think they are often reliable, in an objective sense, and that they can give us knowledge. The natural internal stance of human life assumes that there is a real world, that many questions, both factual and practical, have correct answers, and that there are norms of thought which, if we follow them, will tend to lead us toward the correct answers to those questions. It assumes that to follow those norms is to respond correctly to values or reasons that we apprehend. Mathematics, science, and ethics are built on such norms.

It is difficult to make sense of all this in traditional naturalistic terms. Unless we are prepared to regard most of it as an illusion, this points to a further expansion of our conception of the natural order to include not only the source of phenomenological consciousness—sensation, perception, and emotion—but also the source of our active capacity to think our way beyond those starting points. The question is how to understand mind in its full sense

as a product of nature—or rather, how to understand nature as a system capable of generating mind.

The problem does not arise with respect to the basic forms of perceptual, emotional, and appetitive consciousness that we share **(p.73)** with many other animals. Those mental functions do put us into a complex relation with the world around us, but they seem in principle susceptible of an evolutionary explanation provided it is somehow transformed from the materialist version into something capable of explaining the conscious character of these functions. As I indicated in the last chapter, if such experiences can somehow be added to the evolutionary menu, their roles in enabling creatures to navigate in the world, avoid dangers, find nourishment and shelter, and reproduce all make them potentially adaptive and therefore candidates for natural selection. Perception and desire have to meet certain standards of accuracy to enable creatures to survive in the world: they have to enable us to respond similarly to things that are similar and differently to things that are different, to avoid what is harmful, and to pursue what is beneficial. For most creatures, however, objectivity extends no farther than this. Their lives are lived in the world of appearances, and the idea of a more objective reality has no meaning.

But once we come to recognize the distinction between appearance and reality, and the existence of objective factual or practical truth that goes beyond what perception, appetite, and emotion tell us, the ability of creatures like us to arrive at such truth, or even to think about it, requires explanation. An important aspect of this explanation will be that we have acquired language and the possibilities of interpersonal communication, justification, and criticism that language makes possible. But the explanation of our ability to acquire and use language in these ways presents problems of the same order, for language is one of the most important normatively governed faculties. To acquire a language is in part to acquire a system of concepts that enables us to understand reality.

I am going to set aside at this point all the problems mentioned earlier about the probability of the origin of life and the sufficiency **(p.74)** of random mutation and natural selection to account for the actual evolutionary history of life on earth. The question I want to raise remains even if those problems can be solved for the evolution of plants and lower animals. I will also suppose for the sake of argument that evolutionary theory can be recast in a way that is consistent with antireductionism, so as to make it capable of explaining the appearance of consciousness. The question I now want to

pose is whether our cognitive capacities can be placed in the framework of an evolutionary theory that is in this way no longer exclusively materialist, but that retains the Darwinian structure. It is a hypothetical question, since there may not be such a theory. But I will talk as if there were.

The problem has two aspects. The first concerns the likelihood that the process of natural selection should have generated creatures with the capacity to discover by reason the truth about a reality that extends vastly beyond the initial appearances—as we take ourselves to have done and to continue to do collectively in science, logic, and ethics. Is it credible that selection for fitness in the prehistoric past should have fixed capacities that are effective in theoretical pursuits that were unimaginable at the time? The second problem is the difficulty of understanding naturalistically the faculty of reason that is the essence of these activities. I will begin by considering a possible response to the first problem, before going on to the second, which is particularly intractable.

2

The first problem arises only if one presupposes realism about the subject matter of our thought. We want to know how likely it is, for example, that evolution should have given some human beings **(p.75)** the capacity to discover, and other human beings the capacity to understand, the laws of physics and chemistry. If there is no real, judgment-independent physical world, no judgment-independent truths of mathematics, and no judgment-independent truths of ethics and practical reason, then there is no problem of explaining how we are able to learn about them. On an antirealist view, scientific or moral truth depends on our systematic cognitive or conative responses rather than being something independent to which our responses may or may not conform. The “worlds” in question are all just human constructions. In that case an explanation of how those responses—including our scientific theories—were formed will not have to explain their objective correctness in order to be acceptable (although it will have to explain their internal coherence).

Antirealism of this kind is a more serious option for the moral than for the scientific case. One can intelligibly hold that moral realism is implausible because evolutionary theory is the best current explanation of our faculties, and an evolutionary account cannot be given of how we would be able to discover judgment-independent moral truth, if there were such a thing.¹ But it would be awkward to abandon scientific realism for analogous reasons,

because one would then have to become an antirealist about evolutionary theory as well. This would mean that evolutionary theory is inconsistent with scientific realism and cannot be understood realistically, which seems an excessively strong result. There would be something strange to the point of incoherence about taking scientific naturalism as the ground for antirealism about natural science.

If we leave the assumption of realism in place, the best hope for a naturalistic response to the first problem would be that evolutionary **(p.76)** theory, and in particular evolutionary psychology, is in fact capable of giving a credible account of the success of our cognitive capacities. For factual knowledge, this is the aim of naturalized epistemology. The goal would be to explain how innate mental capacities that were selected for their immediate adaptive value are also capable of generating, through extended cultural evolutionary history, true theories about a law-governed natural order that there was no adaptive need to understand earlier. The evolutionary explanation would have to be indirect, since scientific knowledge had no role in the selection of the capacities that generated it.

The just-so story would go roughly like this. Even in the wild, it isn't just perception and operant conditioning that have survival value. The capacity to generalize from experience and to allow those generalizations, or general expectations, to be confirmed or disconfirmed by subsequent experience is also adaptive. So is a basic disposition to maintain logical consistency in belief, by modifying beliefs when inconsistencies arise. A further, very important step would be the capacity to correct individual appearances not only by reference to other conflicting appearances of one's own but also by reference to how things appear to other perceivers. That requires recognition of other minds, an ability with obvious adaptive potential. The reach of these capacities can be greatly extended and deliberately exercised with the help of language, which also permits knowledge to be collectively created, accumulated, and transmitted. With language we can hold in our minds and share with others alternative possibilities, and decide among them on the basis of their consistency or inconsistency with further observations. Complex scientific theories that entail empirical predictions are therefore extensions of the highly adaptive capacity to learn from experience—our own and that of others.

This story depends heavily on the supposition of a biological origin of the capacity for nonperceptual representation through **(p.77)** language, resulting in the ability to grasp logically complex abstract structures. It is not easy

to say how one might decide whether this could be a manifestation of abilities that have survival value in prehistoric everyday life. In view of the mathematical sophistication of modern physical theories, it seems highly unlikely; but perhaps the claim could be defended.

It is even possible to tell a parallel just-so story about the compatibility between evolutionary theory and moral realism. I am not thinking of the familiar appeal to sociobiology, with its essentially nepotistic interpretation of innate altruistic dispositions. I am not even thinking of the explanation through group selection of dispositions to cooperation in social creatures.² Rather, I have in mind the discovery of general principles of value by rational means analogous to those used elsewhere. Starting from an understanding of innate desires and aversions as immediate impressions of value—of what is good or bad for ourselves or our kin—the discovery of a larger, principle-governed normative domain, or domain of practical reason, in which these immediately apparent values are situated, can again proceed through the capacity to generalize and the disposition to avoid inconsistency.

Generalization would lead to the recognition of value in possible future experiences, in the means to them, and in the lives of creatures other than ourselves. These values are not extra properties of goodness and badness, but just truths such as the following: If something I do will cause another creature to suffer, that counts against doing it. I can come to see that this is true by generalizing from the evident disvalue of my own suffering, and once I recognize the more general truth, my motives will be altered. If there are objective general norms of conduct, this kind of thinking would (p.78) allow us to discover them even if they are no more innate than the laws of physics. As with science, the process of discovery would be impossible without language, interpersonal communication, and cultural memory. In both cases, although the basic capacities employed are adaptive in their simple form, they would permit us to transcend our starting points to discover large domains of truth quite independently of whether such knowledge enhances fitness.

All this is very far-fetched, but no more so than much evolutionary speculation. It requires that mutations and whatever else may be the sources of genotypic variation should generate not only physical structures but phenomenology, desire and aversion, awareness of other minds, symbolic representations, and logical consistency, all having essential roles in the production of behavior. Provided we can assume some global solution to the mind-body problem that allows all this, the rest of the story suggests

that knowledge of objective scientific and moral truth, should there be such things, could result from the exercise of capacities that, in more mundane applications, are at least not inimical to survival. There may not be an insuperable problem of improbability, provided we accept the evolutionary framework itself as probable.

3

However, even if we suppose for the sake of argument that some evolutionary explanation of this kind is true, there is a further problem about thinking of our basic reasoning capacities in this way. It emerges if we contrast the attitude we can reasonably take toward our perceptual and appetitive systems with the attitude we can take toward our reasoning. This will lead to the second problem identified above—the difficulty of understanding reason naturalistically.

(p.79) If we suppose that there is some way to include consciousness in the evolutionary story, then we can understand our visual system, like the visual systems of other species, to have been shaped by natural selection. The specifics of human vision respond to aspects of the world that have been important in the lives of our ancestors. That allows us to continue to rely on the prima facie evidence of our senses, while recognizing that the evidence will sometimes be misleading, selective, or distorted, and that it bears the marks of our particular biological ancestry.

Something similar is possible in our attitudes toward our intuitive judgments of probability, or toward some of our intuitive value judgments (the desire for revenge, for example). We may come to understand those intuitions as rough but useful unreflective responses shaped by natural selection to a fitness-enhancing form in the circumstances in which our forebears lived and died. At the same time, we can recognize that they may need correction or inhibition. Evolutionary self-awareness of this kind is a common feature of our reflective attitudes toward our natural dispositions of hunger, fear, lust, anger, and so forth.

But whenever we take such a reasonable detached attitude toward our innate dispositions, we are implicitly engaged in a form of thought to which we do not at the same time take that detached attitude. When we rely on systems of measurement to correct perception, or probability calculations to correct intuitive expectations, or moral or prudential reasoning to correct instinctive impulses, we take ourselves to be responding to systematic reasons which in themselves justify our conclusions, and which do not get

their authority from their biological origins.³ They *could* not be backed up **(p.80)** in that way. They don't get their authority from their cultural origins, either; on the contrary, the cultural history that has yielded their development is validated as an instance of progress only by the fact that it has led to these methods for increasing the accuracy of our judgments.

Relying on one's vision and relying on one's reason are similar in one respect: in both cases, the reliance is immediate. When I see a tree, I do not infer its existence from my experience any more than I infer the correctness of a logical inference from the fact that I can't help believing the conclusion. However, there is a crucial difference: in the perceptual case I can recognize that I might be mistaken, but on reflection, even if I think of myself as the product of Darwinian natural selection, I am nevertheless justified in believing the evidence of my senses for the most part, because this is consistent with the hypothesis that an accurate representation of the world around me results from senses shaped by evolution to serve that function. That is not a refutation of radical skepticism, since evolutionary theory, like all of science, depends on the evidence of the senses. But it does provide a coherent picture of my place in the world that is consistent with the general reliability of such evidence.

By contrast, in a case of reasoning, if it is basic enough, the only thing to think is that I have grasped the truth directly. I cannot pull back from a logical inference and reconfirm it with the reflection that the reliability of my logical thought processes is consistent with the hypothesis that evolution has selected them for accuracy. That would drastically weaken the logical claim. Furthermore, in the formulation of that explanation, as in the parallel explanation of the reliability of the senses, logical judgments of consistency and inconsistency have to occur without these qualifications, as direct apprehensions of the truth. It is not possible to think, "Reliance on my reason, including my reliance on *this very judgment*, is reasonable **(p.81)** because it is consistent with its having an evolutionary explanation." Therefore any evolutionary account of the place of reason presupposes reason's validity and cannot confirm it without circularity.

Eventually the attempt to understand oneself in evolutionary, naturalistic terms must bottom out in something that is grasped as valid in itself—something without which the evolutionary understanding would not be possible. Thought moves us beyond appearance to something that we cannot regard merely as a biologically based disposition, whose reliability we can determine on other grounds. It is not enough to be able to think that

if there are logical truths, natural selection might very well have given me the capacity to recognize them. That cannot be my ground for trusting my reason, because even that thought implicitly relies on reason in a prior way.

We can suppose that the capacities which enable us to travel far beyond our innate dispositions in representing and responding to the world have appeared in an ancestor and then been preserved in subsequent generations. The appearance of these capacities has to be integrated with the evolutionary process in that they are at least not inimical to fitness, so that they are not extinguished by natural selection. That much seems plausible. But if I am right to think that we can't regard them merely as further instinctive dispositions, some other explanation is needed of what these capacities are.

Just as consciousness cannot be explained as a mere extension or complication of physical evolution, so reason cannot be explained as a mere extension or complication of consciousness. To explain our rationality will require something in addition to what is needed to explain our consciousness and its evidently adaptive forms, something at a different level. Reason can take us beyond the appearances because it has completely general validity, rather than merely local utility. If we have it, we recognize that it can be neither confirmed nor undermined by a theory of its evolutionary origins, nor by any (p.82) other external view of itself. We cannot distance ourselves from it. That was Descartes' insight.

If such a thing appeared on the evolutionary menu, it could have proven its adaptive value locally. Then, with the help of cultural deployment and development, it might have risen to its current position of critical authority, correcting and often overruling the older promptings of perception, instinct, and intuition, and not subject to correction by anything else. Its entrenchment and eventual sovereignty over older instincts is comprehensible—but only if we can understand how such a thing can exist at all.

4

This is the second problem: What is the faculty that enables us to escape from the world of appearance presented by our prereflective innate dispositions, into the world of objective reality? And what, besides consciousness, do we have to add to the biological story to make sense of such a faculty?

The distinctive thing about reason is that it connects us with the truth directly. Perception connects us with the truth only indirectly. When I see a tree, I see it because it is there, but not just because it is there. Perception is not a form of insight: I do not grasp the presence of the tree immediately, even though it may seem so prior to reflection. Rather I am aware of it because the tree causes a mental effect in me in virtue of the character of my visual system, which we may suppose has been shaped by natural selection to react in this way to light reflected from physical objects. Having such a system, together with other perceptual and motivational dispositions, enables me to survive in the world. So it is only in a complicated and indirect sense that when I see a tree, I see it because it is there.

(p.83) But suppose I observe a contradiction among my beliefs and “see” that I must give up at least one of them. (I am driving south in the early morning, and the sun rises on my right.) In that case, I see that the contradictory beliefs cannot all be true, and I see it simply because it is the case. I grasp it directly. It is not adequate to say that, faced with a contradiction, I feel the urgent need to alter my beliefs to escape it, which is explained by the fact that avoiding contradictions, like avoiding snakes and precipices, was fitness-enhancing for my ancestors. That would be an indirect explanation of how the impossibility of the contradiction explains my belief that it cannot be true. But even if some of our ancestors were prey to mere logical phobias and instincts, we have gone beyond that: We reject a contradiction just because we see that it is impossible, and we accept a logical entailment just because we see that it is necessarily true.

In ordinary perception, we are like mechanisms governed by a (roughly) truth-preserving algorithm. But when we reason, we are like a mechanism that can see that the algorithm it follows is truth-preserving. Something has happened that has gotten our minds into immediate contact with the rational order of the world, or at least with the basic elements of that order, which can in turn be used to reach a great deal more. That enables us to possess concepts that display the compatibility or incompatibility of particular beliefs with general hypotheses. We have to start by regarding our prereflective impressions as a partial and perspectival view of the world, but we are then able to use reason and imagination to construct candidates for a larger conception that can contain and account for that part. This applies in the domain of value as well as of fact. The process is highly fallible, but it could not even be attempted without this hard core of self-evidence, on which all less certain reasoning depends. In the criticism and correction of reasoning, the final court of appeal is always reason itself.

(p.84) What this means is that if we hope to include the human mind in the natural order, we have to explain not only consciousness as it enters into perception, emotion, desire, and aversion but also the conscious control of belief and conduct in response to the awareness of reasons—the avoidance of inconsistency, the subsumption of particular cases under general principles, the confirmation or disconfirmation of general principles by particular observations, and so forth. This is what it is to allow oneself to be guided by the objective truth, rather than just by one’s impressions. It is a kind of freedom—the freedom that reflective consciousness gives us from the rule of innate perceptual and motivational dispositions together with conditioning. Rational creatures can step back from these influences and try to make up their own minds. I set aside the question whether this kind of freedom is compatible or incompatible with causal determinism, but it does seem to be something that cannot be given a purely physical analysis and therefore, like the more passive forms of consciousness, cannot be given a purely physical explanation either.

If I decide, when the sun rises on my right, that I must be driving north instead of south, it is because I recognize that my belief that I am driving south is inconsistent with that observation, together with what I know about the direction of rotation of the earth. I abandon the belief because I recognize that it couldn’t be true. If I put money into a retirement account because the future income it generates will be more valuable to me than what I could spend it on now, I act because I see that this makes it a good thing to do. If I oppose the abolition of the inheritance tax, it is because I recognize that the design of property rights should be sensitive not only to autonomy but also to fairness. As the saying goes, I operate in the space of reasons.

The appearance of reason and language in the course of biological history seems, from the point of view of available forms of explanation, something radically emergent—if, as I assume, it cannot be **(p.85)** understood behavioristically. Like consciousness, it presents problems of both constitutive and historical explanation. It appeared long after the emergence of conscious creatures, yet it also seems to be essentially a development of consciousness and ought to be understandable as part of that history. Like consciousness, reason is inseparable from the physical life of organisms that have it, since it acts on the material provided by perception and natural desire and controls action, both directly and indirectly. Any understanding of it will transform our understanding of physical organisms and their development as well.

The great cognitive shift is an expansion of consciousness from the perspectival form contained in the lives of particular creatures to an objective, world-encompassing form that exists both individually and intersubjectively. It was originally a biological evolutionary process, and in our species it has become a collective cultural process as well. Each of our lives is a part of the lengthy process of the universe gradually waking up and becoming aware of itself.

5

This, then, is what a theory of everything has to explain: not only the emergence from a lifeless universe of reproducing organisms and their development by evolution to greater and greater functional complexity; not only the consciousness of some of those organisms and its central role in their lives; but also the development of consciousness into an instrument of transcendence that can grasp objective reality and objective value.

Certain things can be assumed, if there is such a thing as reason. First, there are objective, mind-independent truths of different kinds: factual truths about the natural world, including scientific (p.86) laws; eternal and necessary truths of logic and mathematics; and evaluative and moral truths. Second, by starting from the way things initially appear to us, we can use reason collectively to achieve justified beliefs about some of those objective truths—though some of those beliefs will probably be mistaken. Third, those beliefs in combination can directly influence what we do. Fourth, these processes of discovery and motivation, while mental, are inseparable from physical processes in the organism.

It is trivially true that if there are organisms capable of reason, the possibility of such organisms must have been there from the beginning. But if we believe in a natural order, then something about the world that eventually gave rise to rational beings must explain this possibility. Moreover, to explain not merely the possibility but the actuality of rational beings, the world must have properties that make their appearance not a complete accident: in some way the likelihood must have been latent in the nature of things. So we stand in need of both a constitutive explanation of what rationality might consist in, and a historical explanation of how it arose; and both explanations must be consistent with our being, among other things, physical organisms. The understanding of biological organisms and their evolutionary history would have to expand to accommodate this additional explanatory burden,

as I have argued it must expand beyond materialism to accommodate the explanation of consciousness.

Such an explanation would complete the pursuit of intelligibility by showing how the natural order is disposed to generate beings capable of comprehending it. But the obstacles seem enormous. In light of the remarkable character of reason, it is hard to imagine what a naturalistic explanation of it, either constitutive or historical, could look like.

In the previous chapter I explored the possibility of a reductive account of consciousness, based on some form of universal monism (p.87) or panpsychism. This is modeled on the physical reductionism encouraged by molecular biology, but with an expanded metaphysical basis, in which the physical and the mental are ontologically inseparable. Although it would be a radical departure from the reigning materialist view of nature, the monism required for a reductive but not physically reductionist account of consciousness seems at least conceivable. In answer to the constitutive question, the idea that a complex subject of consciousness might be built up out of minimal protomental elements that are somehow unified simultaneously into an organism and a self has enough potential to merit consideration. Considered as an alternative to an equally speculative emergence of consciousness at high levels of physical organization, it seems relatively credible, in spite of serious problems about the mental part-whole relationship.

However, a reductive account of reason, entirely in terms of the properties of the elementary constituents of which organisms are made, is even more difficult to imagine than a reductive account of consciousness. Rationality, even more than consciousness, seems necessarily a feature of the functioning of the whole conscious subject, and cannot be conceived of, even speculatively, as composed of countless atoms of miniature rationality. The metaphor of the mind as a computer built out of a huge number of transistor-like homunculi will not serve the purpose, because it omits the understanding of the content and grounds of thought and action essential to reason. It could account for behavioral output, but not for understanding. For these reasons, a holistic or emergent answer to the constitutive question comes to seem increasingly more likely than a reductive one as we move up from physical organisms, to consciousness, to reason. This would mean that reason is an irreducible faculty of the kind of fully formed conscious mind that exists in higher animals, and that it cannot be analyzed into (p.88) the

activity of the mind's protomenta parts, in the way that sensation perhaps can be.

But the historical question remains. Even if something entirely new begins to happen when the conscious brain reaches a certain size and level of complexity, an explanation of the existence of that complexity will be adequate only if it also explains the existence of reason as such. (This parallels the demands on an explanation of consciousness as such, discussed in the last chapter.) Suppose we have reason because our brains have reached a level of complexity at which reason emerges. If this is to be an explanation that renders the appearance of reason not a complete accident, it must in some way account not just for the physical complexity itself but for the appearance of just the kind of complexity that is a condition of the emergence of reason. This would not be necessary if one were willing to regard reason as a fluke—a pure side effect of other brain developments. But if that is not acceptable, then an explanation of reason would have to explain the likelihood of the appearance of its biological conditions *qua* conditions of reason, i.e., under that description. The possibilities at this point are too abstractly described to permit any speculation as to whether a reductive causal explanation could do this, but if emergence is the correct answer to the constitutive question about reason, it may be that the historical question will require either a teleological or an intentional solution.

6

I have raised the possibility of teleological principles as part of the natural order in the previous chapter. Teleological explanation may have serious problems, but in this case they are no more serious than those of the alternatives, so the possibility should not be **(p.89)** disregarded. The evolution of mind is part of a single long process of evolutionary descent. It is the latest stage in the evolution of physical organisms, some of which are now governed largely by thought. If we are skeptical about an intentional (theistic) explanation of the existence of reason, and can't make sense of a causal reductionist one, it is natural to speculate that some tendencies in this direction have been at work all along. If physics alone or even a nonmaterialist monism can't account for the later stages of our evolutionary history, we shouldn't assume that it can account for the earlier stages. Indeed, when we go back far enough, to the origin of life—of self-replicating systems capable of supporting evolution by natural selection—those actually engaged in research in the subject recognize that they are very far from even formulating a viable explanatory hypothesis of the traditional

materialist kind. Yet they assume that there must be such an explanation, since life cannot have arisen purely by chance.⁴

In fact, that assumption may be based on a confusion. In an important paper, Roger White has argued that the search for an explanation of the origin of life in terms of the nonpurposive principles of physics and chemistry—an explanation that will reveal that the origin of life is not merely a matter of chance but something to be expected, or at least not surprising—is probably motivated by the sense that life can't be a matter of chance because it looks so much as though it is the product of intentional design. But the hypothesis of intentional design is ruled out as unscientific. So it seems natural to conclude that the only way left for life not to be a matter of chance is for it somehow to be made likely by physical law. **(p.90)** But as White points out, this inference is illegitimate. Here is what he says:

The line of reasoning ... is something like the following. That molecular replicating systems appear to be designed by an agent is sufficient to convince us that they didn't arise by chance. But in scientific reasoning, non-intentional explanations are to be preferred, if possible (some would say at all costs), to intentional ones—hence the motivation to find a non-intentional explanation of life.

It should be clear however, that even granting the appropriateness of a preference for non-intentional explanations, this line of reasoning is confused. In general, if B_I [the hypothesis that the process that led to S was intentionally biased] raises the likelihood of S , then S confirms B_I to at least some degree, and may thereby disconfirm C [the chance hypothesis]. But it does not follow that S confirms B_N [the hypothesis that the process was non-intentionally biased] one iota. S confirms B_N *only if* B_N raises the likelihood of S . If the reason we doubt the Chance Hypothesis is that we suspect that life is due in part to intelligent agency, this by itself gives us no reason to expect there to be a non-intentional explanation for life. If on reflection we do not find the hypothesis of intentional biasing acceptable, then we are left with no reason at all to doubt that life arose by chance.⁵

Much of White's paper is taken up with arguing that life is no more to be expected on the assumption of B_N —the hypothesis of nonintentional bias—than on the assumption of chance. That is because **(p.91)** even if there

is nonintentional bias toward certain outcomes resulting from purposeless physical law, it could be a bias toward any type of outcome whatever, so it cannot make the appearance of life more likely than anything else. As White says,

What makes certain molecular configurations stand out from the multitude of possibilities seems to be that they are capable of developing into something that strikes us as rather marvelous, namely a world of living creatures. But there is no conceivable reason that blind forces of nature or physical attributes should be biased toward the marvelous.⁶

By contrast, intentional bias is limited as a hypothesis by some rough assumptions about the motives that give rise to intentions. (Thus one cannot claim about just any outcome *S*, however random or arbitrary, that it gives evidence that the process that led to it was intentionally biased, simply on the ground that it is rendered likely by the hypothesis that it was produced by a being who wanted precisely *S* to occur. Any argument from design depends on more restrictive general assumptions about what kinds of things a designer might want to produce.)

I am drawn to a fourth alternative, natural teleology, or teleological bias, as an account of the existence of the biological possibilities on which natural selection can operate. I believe that teleology is a naturalistic alternative that is distinct from all three of the other candidate explanations: chance, creationism, and directionless physical law. To avoid the mistake that White finds in the hypothesis of nonintentional bias, teleology would have to be restrictive in what it makes likely, but without depending on intentions or motives. This would probably have to involve some conception of an increase in value through the (p.92) expanded possibilities provided by the higher forms of organization toward which nature tends: not just any outcome could qualify as a *telos*. That would make value an explanatory end, but not one that is realized through the purposes or intentions of an agent. Teleology means that in addition to physical law of the familiar kind, there are other laws of nature that are “biased toward the marvelous.”

The idea of teleology as part of the natural order flies in the teeth of the authoritative form of explanation that has defined science since the revolution of the seventeenth century. Teleology would mean that some natural laws, unlike all the basic scientific laws discovered so far, are temporally historical in their operation. The laws of physics are all equations specifying universal relations that hold at every time and place

among mathematically specifiable quantities like force, mass, charge, distance, and velocity. In a nonteleological system the explanation of any temporally extended process has to consist in the explanation, by reference to those laws, of how each state of the universe evolved from its immediate predecessor. Teleology, by contrast, would admit irreducible principles governing temporally extended development.

The teleology I want to consider would be an explanation not only of the appearance of physical organisms but of the development of consciousness and ultimately of reason in those organisms. But its form can be described even if we stay at the physical level. Natural teleology would require two things. First, that the nonteleological and timeless laws of physics—those governing the ultimate elements of the physical universe, whatever they are—are not fully deterministic. Given the physical state of the universe at any moment, the laws of physics would have to leave open a range of alternative successor states, presumably with a probability distribution over them.

Second, among those possible futures there will be some that are more eligible than others as possible steps on the way to the formation of more complex systems, and ultimately of the kinds of **(p.93)** replicating systems characteristic of life. The existence of teleology requires that successor states in this subset have a significantly higher probability than is entailed by the laws of physics alone—simply because they are on the path toward a certain outcome. Teleological laws would assign higher probability to steps on paths in state space that have a higher “velocity” toward certain outcomes.⁷ They would be laws of the self-organization of matter, essentially—or of whatever is more basic than matter.

This is a frankly teleological hypothesis because the preferred transitions do not have a higher probability in virtue of their intrinsic immediate characteristics, but only in virtue of temporally extended developments of which they form a potential part. In other words, some laws of nature would apply directly to the relation between the present and the future, rather than specifying instantaneous functions that hold at all times. A naturalistic teleology would mean that organizational and developmental principles of this kind are an irreducible part of the natural order, and not the result of intentional or purposive influence by anyone. I am not confident that this Aristotelian idea of teleology without intention makes sense, but I do not at the moment see why it doesn't.

What are the alternatives? Any alternative must include the constitutive possibility, in the character of the elements of which the world is composed, of their combination into living organisms with the properties of consciousness, action, and cognition which we know they have. But given this possibility, the historical question of why such **(p.94)** organisms arose could in principle receive two very different nonteleological answers. First, there is the hypothesis that the initial appearance of a code-governed replicating system that started the evolutionary process was a cosmic accident, and that subsequent accidental mutations provided the set of successive candidates on which natural selection operated to generate the history of life. This hypothesis makes the outcome too accidental to count as a genuine explanation of the existence of conscious, thinking beings as such.

Second, for theists there is the intentional alternative: divine intervention to create life out of the basic material of the world, and perhaps also to guide the process of evolution by natural selection, through the intentional production and preservation of some of the mutations on which natural selection operates along the way.⁸ This could be combined with either a reductive or an emergent answer to the constitutive question. A creationist explanation of the existence of life is the biological analogue of dualism in the philosophy of mind. It pushes teleology outside of the natural order, into the intentions of the creator—working with completely directionless materials whose properties nevertheless underlie both the mental and the physical. If God at some point in the past constructed DNA or one of its predecessors out of its elements, that dispenses with the need for any explanation of the capacity of the elements to organize themselves in this apparently purposive way.

That would require only that the existence of DNA be a physical possibility—in chemical space, so to speak. And if we extend the case to consciousness and reason, it would require that conscious and rational subjects supported by brains of the right kind be mental possibilities. But in the creationist picture, the natural order accounts for the physical possibility of DNA in the same way that it accounts for the physical possibility of an airplane or a telephone or a computer. **(p.95)** Those possibilities are all explained by physics alone: it is only their actualization that involves a designer, and something analogous would be true for animal consciousness—a surprising way in which the protopsychic elements of the world can be combined. So biological and mental organization are no more part of the natural order in

the creationist view than airplanes or telephones are. The laws of nature entail their possibility, but they do not explain their actuality.

My preference for an immanent, natural explanation is congruent with my atheism. But even a theist who believes God is ultimately responsible for the appearance of conscious life could maintain that this happens as part of a natural order that is created by God, but that it does not require further divine intervention. A theist not committed to dualism in the philosophy of mind could suppose that the natural possibility of conscious organisms resides already in the character of the elements out of which those organisms are composed, perhaps supplemented by laws of psychophysical emergence. To make the possibility of conscious life a consequence of the natural order created by God while ascribing its actuality to subsequent divine intervention would then seem an arbitrary complication. Some form of teleological naturalism should for these reasons seem no less credible than an interventionist explanation, even to those who believe that God is ultimately responsible for everything.⁹ (p.96)

Notes:

(1) See Sharon Street, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127 (2006): 109–66. I pursue this topic in the next chapter.

(2) See Elliott Sober and David Sloan Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge, MA: Harvard University Press, 1998).

(3) See the last chapter of my *The Last Word* (New York: Oxford University Press, 1997); I am here continuing the discussion of questions posed there.

(4) See note 8 of chapter 1 for relevant references, and also Iris Fry, *The Emergence of Life on Earth: A Historical and Scientific Overview* (New Brunswick, NJ: Rutgers University Press, 2000).

(5) Roger White, "Does Origins of Life Research Rest on a Mistake?" *Noûs* 41 (2003), 475.

(6) *Ibid.*, 467.

(7) See Hawthorne and Nolan, "What Would Teleological Causation Be?"

(8) See Alvin Plantinga, *Where the Conflict Really Lies: Science, Religion, and Naturalism* (New York: Oxford University Press, 2011).

(9) Are there any other alternatives? Well, there is the hypothesis that this universe is not unique, but that all possible universes exist, and we find ourselves, not surprisingly, in one that contains life. But that is a cop-out, which dispenses with the attempt to explain anything. And without the hypothesis of multiple universes, the observation that if life hadn't come into existence we wouldn't be here has no significance. One doesn't show that something doesn't require explanation by pointing out that it is a condition of one's existence. If I ask for an explanation of the fact that the air pressure in the transcontinental jet is close to that at sea level, it is no answer to point out that if it weren't, I'd be dead. Compare Leslie, *Infinite Minds*, 207. See also Roger White, "Fine-Tuning and Multiple Universes," *Noûs* 34 (2000): 260-76; and Ian Hacking, "The Inverse Gambler's Fallacy: The Argument from Design. The Anthropic Principle Applied to Wheeler Universes," *Mind* 96 (1987): 331-40, for further problems with appealing to multiple universes to explain the character of this one.

