View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Asbury Theological Seminary

Faith and Philosophy: Journal of the Society of Christian Philosophers

Volume 19 | Issue 1 Article 1

1-1-2002

Irreducible Complexity and Darwinian Gradualism: A Reply to Michael J. Behe

Paul Draper

Follow this and additional works at: https://place.asburyseminary.edu/faithandphilosophy

Recommended Citation

Draper, Paul (2002) "Irreducible Complexity and Darwinian Gradualism: A Reply to Michael J. Behe," *Faith and Philosophy: Journal of the Society of Christian Philosophers*: Vol. 19 : Iss. 1 , Article 1.
Available at: https://place.asburyseminary.edu/faithandphilosophy/vol19/iss1/1

This Article is brought to you for free and open access by the Journals at ePLACE: preserving, learning, and creative exchange. It has been accepted for inclusion in Faith and Philosophy: Journal of the Society of Christian Philosophers by an authorized editor of ePLACE: preserving, learning, and creative exchange.

IRREDUCIBLE COMPLEXITY AND DARWINIAN GRADUALISM: A REPLY TO MICHAEL J. BEHE

Paul Draper

In *Darwin's Black Box*, Michael J. Behe argues that, because certain biochemical systems are both irreducibly complex and very complex, it is extremely unlikely that they evolved gradually by Darwinian mechanisms, and so extremely likely that they were intelligently designed. I begin this paper by explaining Behe's argument and defending it against the very common but clearly mistaken charge that it is just a rehash of William Paley's design argument. Then I critically discuss a number of more serious objections to the argument. I conclude that, while Behe successfully rules out some Darwinian paths to the biochemical systems he discusses, others remain open. Thus, his argument against Darwinian gradualism (and ipso facto his argument for intelligent design) is at best incomplete.

In *Darwin's Black Box*, Michael J. Behe[1] argues that, because certain biochemical systems are both irreducibly complex and very complex, it is extremely likely they were designed by an intelligent agent. Behe, who is a biochemist, wrote this book for a popular audience, and it has received a great deal of attention, much of it quite positive, from the popular press.[2] Not surprisingly, it has also been ridiculed by many in the scientific community. Yet it has managed to impress a number of very good philosophers,[3] not, I suspect, because of its philosophical content, which is very limited, but rather because it contains a new and interesting challenge to Darwinism, which is arguably the most philosophically significant scientific theory ever formulated. My goal in this paper will be to show that, while this challenge is both more original and, with a few modifications, more powerful than many of Behe's critics realize, it is incomplete and for that reason does not refute Darwinism.

I will divide the body of my paper into three main sections, followed by a brief conclusion. In the first section, I will explain Behe's design argument. In the second, I will defend it against the very common but clearly mistaken charge that it is just a rehash of William Paley's design argument, which Darwin refuted over a century ago. And in the third section, I will critically discuss a number of more serious objections to Behe's argument. Some of these objections can be answered by modifying his argument. But others expose sizable holes in the argument, and I am unable to see how Behe could fill them.

FAITH AND PHILOSOPHY
Vol. 19 No. 1 January 2002

All rights reserved

I

Behe's design argument has two stages. The first stage, which is developed in the first two parts of his tripartite book, appeals to the existence of biochemical "machines" in an attempt to show that Darwinism is very probably false. It is important to be clear, however, on exactly what aspect of Darwinism Behe is attacking. He disputes neither common ancestry nor the operation of natural selection. Rather, it is Darwin's gradualism that Behe rejects. Specifically, he attempts to meet a challenge that Darwin himself issued: "If it could be demonstrated that any complex organ existed, which could not possibly have been formed by *numerous, successive, slight modifications*, my theory would absolutely break down."[4] Of course, there is an important connection between natural selection and gradualism. Natural selection can work on small successive changes or on sudden big changes, but, as Richard Dawkins has pointed out,[5] it cannot *explain* a complex system like the human eye unless it operates on a long series of small changes. This is why Darwin says his theory would break down if it could be shown that a complex organ could not have evolved gradually. If such an organ appears suddenly, then it does not explain that organ's complexity to point out that the organ gave the organism in which it first appeared an advantage in the struggle to survive and so through heredity became increasingly common in the population to which that organism belonged. So by attacking gradualism, Behe is attacking Darwin's view that natural selection can explain biological complexity.

The second stage of Behe's argument, which is defended in the third part of his book, attempts to show that, if certain biochemical systems were not put together gradually, then it is highly probable they were designed by one or more intelligent agents. This part of Behe's argument is not sufficiently developed to be of interest to philosophers. Behe, keep in mind, is a biochemist, and so his failure to discuss the logical details of his inference from the falsity of Darwinian gradualism to the truth of intelligent design is not surprising. He believes that the correctness of this inference is or should be obvious (at least once complexity theory and symbiosis are ruled out as alternative non-gradualist explanations) and for that reason does little (by philosophical standards) to defend it.[6] Thus, I will ignore this part of Behe's argument, focusing instead on his very interesting challenge to Darwinian gradualism.

Behe bases this challenge on the claim that certain biochemical systems like the cilium and the blood clotting cascade are *irreducibly complex*–that is, they are "composed of several well-matched, interacting parts that contribute to the basic function, wherein the removal of any of the parts causes the system to effectively cease functioning" (p. 39). To help us understand this definition, Behe provides a simple non-biological example: a mousetrap of the common household variety. This system fits his definition nicely. It is composed of several–five to be exact–interacting, well-matched parts, all of which contribute to the basic function of the system, which is to kill mice. Most importantly, the removal of any of the parts would cause the system to cease functioning: a hammer is needed to deliver the killing blow, a spring and holding bar are required to arm the hammer, a catch is

crucial both for arming the hammer and for detecting the pressure of small feet, and a platform is essential for anchoring the other parts.

How, according to Behe, does the existence of irreducibly complex biochemical systems show that Darwinian gradualism is false? His argument–as stated in the first part of his book–proceeds as follows. If gradualism is true, then these biochemical systems are produced either directly or indirectly by slight, successive modifications of a precursor system. But they "cannot be produced directly (that is, by continuously improving the initial function, which continues to work by the same mechanism) by slight, successive modifications of a precursor system, because any precursor to an *irreducibly* complex system that is missing even a single part is by definition nonfunctional" (p. 39; my emphasis). And while it is possible that they are gradually produced indirectly–that is, via an evolutionary path on which a change occurs either in the system's function or in the mechanism by which the system performs its function–this is extremely improbable because they are *very* complex: "As the complexity of an interacting system increases, . . . the likelihood of . . . an indirect route drops precipitously" (p. 40). Behe concludes that gradualism is very probably false–it is highly unlikely that these biochemical systems are produced by slight, successive modifications of a precursor system.

Notice that, in this initial statement of the argument, Behe uses the *irreducibility* of a system's complexity to rule out (as conceptually impossible) direct evolutionary routes–that is, routes that do not involve a change in function or mechanism–while the fact that a system is *very* complex–it has very many parts (p. 48)–is used to rule out (as highly improbable) indirect evolutionary routes. Many of Behe's critics ignore this dual aspect of Behe's argument when they offer alleged counterexamples to Behe's position that Darwinian mechanisms did not produce the sort of systems he discusses. Either they give examples of very complex systems that evolved gradually but are not (at least so far as we can tell) irreducibly complex, or they give examples of irreducibly complex systems that evolved gradually, but ignore the fact that those systems do not exhibit a high degree of complexity.

An example of the first sort of mistake can be found in a reply to Behe offered by Niall Shanks (a philosopher) and Karl H. Joplin (a biologist).[7] They argue that complex biochemical systems typically exhibit considerable redundancy, which implies that parts can be removed without a complete loss of function. Such redundancy also helps us to see how such systems could have evolved step by step. But Behe is not committed to claiming that all biochemical systems exhibit irreducible complexity or that no biochemical systems evolved gradually in a Darwinian fashion. And Shanks and Joplin offer no evidence that the biochemical systems to which Behe appeals exhibit redundant complexity. Thus, no damage is done to Behe's argument when they point out that other very complex biochemical systems do exhibit redundant and hence reducible complexity and suggest how those systems might have evolved gradually.[8]

An example of the second sort of mistake can be found in a recent book by Kenneth R. Miller (a cell biologist), who tells the following well-known Darwinian success story:

> The three smallest bones in the human body–the *malleus*, *incus*, and *stapes*–carry sound vibrations across the middle ear, from the membrane-like tympanum (the eardrum) to the oval window. This five-component system perfectly fits the criterion of irreducible complexity–if any one of its parts is taken away or modified, hearing is lost. This is the kind of system that evolution supposedly cannot produce, since, as Behe has said, it would have to "arise as an integrated unit, in one fell swoop." Unfortunately for design theorists, the fossil record elegantly and precisely documents exactly how this system formed.
>
> During the evolution of mammals, over several million years, two of the bones that originally formed the rear portion of the reptilian lower jaw were gradually pushed backwards and reduced in size until they migrated into the middle ear, forming the bony connections that carry vibrations into the inner ears of present-day mammals. This is an example of a system of perfectly formed interlocking components, specified by multiple genes, that was gradually refashioned and adapted for another purpose altogether–something that evolution's critics [including, Miller believes, Behe] claim to be *impossible*.[9]

Miller concludes: "Remember Behe's statement that 'any precursor to an irreducibly complex system that is missing a part is by definition nonfunctional'? Well, there's just no other word for it–that statement is *wrong*."[10]

But Miller's example of the mammalian ear certainly does not justify this conclusion. For Behe does not mean to rule out *by definition* functioning precursors to irreducibly complex systems that are missing parts but serve either a different function or the same function by a different mechanism. And if Miller is right that the mammalian ear is irreducibly complex,[11] then that is precisely what we have in the case of the malleus, incus, and stapes. We have a perfect example of how it is possible, as Behe admits, for an irreducibly complex system to evolve gradually by an *indirect* route. One of the three bones of the mammalian middle ear came from a reptilian ear, which functioned by a different mechanism. The other two came from a reptilian jaw, the function of which was obviously not to transmit sound vibrations.

Thus, if this example is any threat at all to Behe's argument, it is because it provides evidence against Behe's claim that the indirect, gradual evolution of the biological systems he discusses is very unlikely. But it fails even to do this. For even if the mammalian ear is irreducibly complex, it is not the irreducible complexity of the systems Behe discusses that is supposed to make their gradual indirect evolution unlikely, but rather their high degree of complexity. And the mammalian ear has only five components. Thus, though it may be *irreducibly* complex, it is not *very* complex, and hence proof of its gradual indirect evolution does no damage to Behe's argument.[12]

II

Behe's critics have raised three main sorts of objections to his argument against Darwinian gradualism. Some deny that the biochemical systems upon which Behe bases his case are irreducibly complex. Some maintain

that the indirect Darwinian evolution of these systems is not unlikely. And some attack Behe's position on direct Darwinian routes–they claim that, even without any change in function or mechanism, the gradual evolution of an irreducibly complex system is possible. I will raise or discuss objections of all these sorts in Section III. But first, I would like to defend Behe against what is perhaps the most common criticism of his book.

Most reviewers complain that the book is unoriginal–that Behe has just taken William Paley's old design argument and dressed it up in new biochemical clothing. For example, Peter Atkins (a chemist) claims that Behe's 'book is a tiresome reworking at the molecular level of the timeworn "design argument."'[13] And Miller says that "Behe would like us to believe that he has discovered a new biological principle called 'design.' But the real news in *Darwin's Black Box* is not design. . . . It is, instead, that a classic argument from the nineteenth century could be attractively rewrapped in the shiny packaging of biochemistry."[14] Miller adds that Behe's argument is doomed to fail, on the grounds that Paley's argument was refuted by Darwin (and more recently by Dawkins).[15] Both Atkins and Miller are confident that what did not work at the level of the organism will not work at the level of the cell.[16] Behe himself appears to lend credibility to this charge by emphasizing the similarity of his argument to Paley's and by contending that Paley's argument has never been refuted (p. 213). I believe, however, that this charge is both misleading and unfair, especially when used to justify pessimism about the success of Behe's argument. Though the truth of my belief may already be obvious to anyone who has read both Paley's *Natural Theology* and the first section of this paper, it is nevertheless useful to compare Paley's argument to Behe's, both in order to highlight certain distinctive features of Behe's argument and in order to explain why Behe makes the provocative claim that Paley has never been refuted.

Paley was even less interested in the logical details of his inference from natural mechanical order to design than Behe is. He used an analogy to defend his inference (his famous watch example[17]), but obviously that's no reason to think he intended his argument to be analogical. If we interpret it instead as an instance of inference to the best available explanation, then it can be summarized as follows:

1. Some natural systems (e.g. the human eye) are mechanically ordered (i.e. they exhibit the same sort of order as watches and other machines produced by human beings).
2. Intelligent design is a very good explanation of mechanical order.
3. No other explanation (or no equally good explanation) of mechanical order is available.
4. Every instance of mechanical order has an explanation.

So,

5. Some natural systems were (probably) intelligently designed.

According to Elliott Sober, this argument was not refuted until Darwin published his *Origin of Species*. Sober claims that, although Hume may have refuted Paley's argument, interpreted as an argument by analogy, Hume

offered no alternative explanation of biological order, and so did not refute Paley's argument, interpreted as an inference to the best available explanation. That refutation was not provided until Darwin showed that premise (3) is false by providing a better explanation of mechanically ordered biological systems than intelligent design.[18] Behe agrees with Sober that Hume failed to refute Paley's argument (pp. 217-219), but denies that Darwin did any better. Because mechanical order includes irreducible complexity, Behe's argument against gradualism proves, in his opinion, that Darwinian explanations of very complex natural machines do not work.[19]

But Behe's contribution to Paley's case for design is not limited to defending premise (3). Behe holds that Paley lacked clear examples of natural machines. Paley's favorite examples were large biological structures like the eye. Even if we ignore the (alleged) problem that such systems are typically not discrete systems but rather contain a number of distinct subsystems, each with its own special function (p. 38), Behe identifies a further (in my opinion more serious) problem, namely, that the working parts of large biological systems like the eye typically include a very large number of different types of molecules, and we cannot even list all of these molecules let alone specify how they contribute to the function of such systems. It follows that we don't know whether or not such systems are irreducibly complex (i.e. whether or not they are "machines" in the relevant sense) and so, Behe concludes, don't know whether or not they evolved gradually (p. 41).

Some scientists have objected that a mousetrap is also composed of molecules, and yet "to see that a mousetrap is irreducibly complex, we don't have to work out its chemistry."[20] But those molecules are not *working* parts of the mousetrap. Watchmakers at Paley's time, who had no idea what a molecule is,[21] could list every single working part of a watch and explain how each contributes to the watch's function of keeping time, whereas biologists cannot do the same for the eye or its subsystems, even today. So Paley was in a position to know that a watch is irreducibly complex, but we still don't know whether or not the eye or its subsystems are. This is why biochemistry is essential to Behe's argument. According to Behe, we now know that a number of discrete molecular systems are both very complex and irreducibly complex: biochemists have identified the specific functions of these systems, they can list all of their many well matched parts, and they know that those parts are required for the systems to function (p. 41). This could not have been done when Paley wrote or even when Darwin wrote, because, prior to the second half of the 20th Century, the cell was still a "black box"–no one even knew what its working parts were, let alone how they worked.

It is clear, then, that Behe has done more than just "spiff up" Paley's design argument with some new molecular examples of biological order. It is Behe's position that, while Paley has never been refuted–neither his premises nor his inference have been shown to be defective–he still needs help on two fronts. He needs a refutation of Darwinian explanations of natural machines to show that his third premise is true, and he needs some proven examples of natural machines to establish that his first premise is true. Whether or not Behe's attempt to provide such help is ultimately successful, it is certainly original. For he emphasizes in a way Paley never did both the idea that mechanical order includes irreducible complexity and

the importance of this idea for Paley's inference from mechanical order to design. Moreover, Behe attempts to refute Darwinian gradualism, which is, of course, an opponent Paley never had to face. And while a large part of Behe's book does involve presenting new biochemical examples of mechanical order, this is accompanied by an argument for the conclusion that such examples are needed–that Paley's examples did not clearly demonstrate the existence of natural machines. Given all this, Behe's book should certainly not be dismissed as unoriginal. It deserves to be carefully evaluated.

III

A. *Are Behe's Systems Irreducibly Complex?*

A number of biochemists and biologists, and even one or two philosophers, challenge Behe's claim that the systems he discusses in the second part of his book are irreducibly complex. Unfortunately, some of these challengers insist on *defining* "irreducible complexity" as "complexity that cannot (or at least did not) evolve gradually." On this definition, Behe's claim that some biological systems are irreducibly complex becomes the conclusion of his argument against gradualism rather than a premise. But some of those who claim that Behe's systems are not irreducibly complex are careful to stick to Behe's definition–they really do mean to be claiming that these systems can function (though perhaps not as well) without some of their parts. For example, Miller[22] challenges Behe's claim that the sort of cilium discussed by Behe is irreducibly complex by pointing out that comparative studies on a wide variety of organisms show that the common "textbook" claim that the only working cilia have nine pairs of microtubules surrounding two central microtubules is false. In fact, some function with only three outer pairs of microtubules and no central microtubules.[23] David W. Ussery (a biochemist) points out that, while Behe claims that the bacterial flagellum requires over forty proteins to function, some bacteria have functioning flagella composed of only 33 proteins, and these fully functional simpler flagella may be able to function well enough to be useful to a bacterium with still fewer parts. Ussery also challenges Behe's claim that certain cellular transport systems are irreducibly complex, citing as evidence the much simpler transport systems found in bacterial cells.[24] Finally, one does not have to be a biologist to realize that the antibody-diversity system, which is one of three subsystems of the immune system that Behe claims to be irreducibly complex, would not "effectively cease to function" if a few antibody molecules were removed from it.[25]

Additional support for this objection comes (unintentionally) from Behe himself. For a careful reading of the second part of his book reveals that he rarely even claims let alone proves that all of the parts of the systems he discusses are required for those systems to function. His first two examples are the cilium and the flagellum, which he says are systems of "staggering complexity, with dozens or even hundreds of precisely tailored parts" (p. 73). But are all of these parts necessary for cilia and flagella to function? In the case of the cilium, Behe makes no effort to defend an affirmative answer to this question. Instead, he argues for the necessity of

three parts (microtubules for paddling, dynein arms to serve as motors, and nexin arms to transmit the force of the motors to the paddles), and then claims that "It is *very likely* that *many* of the parts we have not considered here are required for any cilium to function in a cell" (p. 73; my emphases). In the case of the bacterial flagellum, he does claim that it requires more than forty proteins to function, but he does not defend that claim and, as we shall see, implicitly admits that it is false in a later paper.

Behe's third example of an irreducibly complex system is the vertebrate blood clotting system. Behe specifies the system's components (over two dozen proteins), but instead of showing that they are all required for the system to function, Behe admits that the details of much of this very complex system are not very well known. Indeed, there is a fork in the pathway (traced backwards), which suggests that part of the system exhibits redundant and hence reducible complexity. And though he says that "the *entire* blood clotting system is irreducibly complex" (p. 87; my emphasis), what he actually argues for is that the four components of the part of the system that is well known–the part after the fork in the pathway–are all required for the system to work (p. 86). He also claims that each step in the pathway is irreducibly complex, since each step requires both a proenzyme and an activating enzyme to switch on the proenzyme (p. 87).

Behe's four other examples also fall short of the expectations created by his definition of "irreducible complexity." He claims that transporting proteins from one part of the cell to another for disposal from the cell involves the services of dozens of proteins (pp. 106-108), but then says that less than ten components are necessary (pp. 109-110). He also claims that three subsystems of the immune system–clonal selection, antibody diversity, and the complement system–are each irreducibly complex.[26] But once again, Behe does not even try to show that all of the functioning parts of these systems are necessary. Instead, though he wouldn't put it this way, he seems to defend the position that these systems in some sense *contain* irreducible complexity. For example, he points out that the complement system is a cascade that, like the blood-clotting cascade, includes steps that require each of two components to function.

The bottom line is that Behe doesn't deliver in the second part of his book what he promised in the first part. Both his definition of "irreducible complexity" and his example of the five part mousetrap suggest that, in an irreducibly complex system, *all* of the interacting parts that contribute to the system performing its function are required for that function. This is also spelled out very clearly when he describes the two steps for determining if a system is irreducibly complex. The first step, he says, is to identify the system's function and its components–i.e. the parts of the system that contribute to its function. "The second step . . . is to ask if *all* of the components are required for the function" (p. 42; my emphasis). Thus, Behe fails to live up to his own standards for determining that a system is irreducibly complex. The systems upon which he bases his case contain parts that contribute to the system's function, yet either are not essential for that function or at least have not been shown by Behe to be so.

In a paper published after the book, Behe in effect changes his definition of "irreducible complexity" so as to avoid this particular problem. In

response to Ussery's claim that flagella can function with fewer than forty proteins, he says that "some systems may have parts that are necessary for a function, plus other parts that, while useful, are not absolutely required. Although one can remove the radio from a car and the car will still work, one can't remove the battery or some other parts and have a working car."[27] His example here is a bit off the mark, since a car's radio does not contribute to a car's (most obvious) function and so is not a "component" of the system in the relevant sense. But replace "radio" with "tires" and the example makes the point he wants. Although not all of the parts of a car that contribute to its function are required for that function, many are, and for that reason the car qualifies as an irreducibly complex system. Notice that, by responding to Ussery in this way, Behe in effect admits that his original claim that forty proteins are required for a bacterial flagellum to function is false. But more importantly, this response presupposes a new definition of "irreducible complexity," one that implies that a system is irreducibly complex even if it has working parts that are not essential for it to function, so long as it has (at least) two interacting and closely matched parts which are essential. Given this definition, it would seem that the biochemical systems Behe discusses are indeed irreducibly complex.

This change in definition, however, does not solve the real problem here, which is that, while Behe (on the new definition) has provided examples of systems that are both irreducibly complex and very complex, he has not proven that these systems are irreducibly very complex. This is a problem because great complexity that is reducible is not an insurmountable obstacle for direct Darwinian evolution, and the irreducible complexity of a relatively simple system (like the mammalian ear) is not an insurmountable obstacle for indirect Darwinian evolution. Thus, even if Behe is right that a direct Darwinian path cannot produce irreducible complexity and an indirect Darwinian path is very unlikely to produce a high degree of complexity, this leaves open the very real possibility of an indirect Darwinian path producing an irreducibly complex minimal system and then a direct path adding to that system all of the extra complexity that improves the system's function but is not necessary for it.

I do not know how Behe would respond to this new objection. Perhaps he could show that some of the systems he discusses have sufficiently many required components to rule out as unlikely an indirect route's producing the minimal system. Or perhaps he could appeal to other reasons besides great complexity to rule out as unlikely the production of the minimal system by indirect evolution. This second sort of reply, it might be argued, is already implicit in Behe's book. For although his argument as stated in the first part of his book appeals only to great complexity in order to rule out indirect routes, we shall see in the next section that, when he actually applies his reasoning to the various systems he discusses in the second part of the book, great complexity (in the sense of a multiplicity of parts) is not the only nor even the chief reason he offers for rejecting indirect routes. Unfortunately for Behe, however, we shall also see that his chief reason–the specificity of the parts of the system–does not rule out every sort of indirect route.[28]

B. *Are Indirect Routes Improbable?*

Recall that indirect evolutionary routes are ones involving a change in function or a change in mechanism. The simpler a system, the more likely it is that indirect evolution by itself could account for that system. Behe sometimes seems to forget this fact. Consider, for example, his argument for the conclusion that a new two-part step in the blood clotting cascade could not be added in a Darwinian fashion to a simpler cascade. He correctly points out that the probability of both parts appearing even close to simultaneously is negligibly low. Thus, realistically, one of the parts would have to be introduced first. But then, in an effort to reach the conclusion that the part that appears first would disappear long before the arrival of the other part, he claims that at best the first part would do nothing and so would not be selected (p. 87). This claim, however, fails to establish the desired conclusion for two reasons. First, even if it is true, the first part might appear coincidentally in an organism that is for other reasons the fittest and thus might persist despite doing nothing.[29] Second, Behe fails to show that the claim is true, that the first part would do nothing. For from the fact that, for example, a new inactive proenzyme would not by itself contribute to blood clotting, it doesn't follow that it would do nothing. It might very well perform some other function, resulting in its selection. Then, when the necessary activating enzyme appeared, both could be "recruited" as a new step in the blood clotting cascade. (Actually, given gene duplication, for which much evidence exists in the case of the blood clotting cascade, the performance of new functions does not imply that old functions cease to be performed, making such indirect paths even easier to travel.)

Of course, Behe may have good reasons based on biological details of which I am ignorant for believing that such indirect routes are very unlikely in this specific case. But he does not tell us what these reasons are. Instead, he often challenges his critics to explain in detail exactly how the system in question evolved. I certainly cannot do that–indeed, maybe no one can at this time. But I do not claim to be able to prove that Darwinism explains the development of these systems. Rather, it is Behe who claims he can show that no such explanation exists. And obviously the fact that no one can specify all of the specific steps in the evolution of these systems does not show that the sort of indirect route discussed above did not produce them.

Another worry I have about Behe's analysis of indirect routes is that, while there are many different sorts of indirect Darwinian routes, Behe discusses only one at any length. I will begin by examining what he says about that one, and then show that his reasons for rejecting it fail to rule out a different sort of indirect route. The type of indirect path to which he gives most of his attention is one on which the required parts of an irreducibly complex system evolve separately for different purposes but then unite in the service of a single new function. Keep in mind that he admits such routes are possible. A system S which performs some function F might gradually evolve despite having more than one required component if those components evolve separately (at different times), each serving some distinct function other than F, and then, when all of the parts of S are present, they begin to interact with each other, forsaking their old functions

(which might not be as beneficial as F or might continue to be performed because of some sort of redundancy) for the sake of F.

Behe holds that such paths are very unlikely to account for the systems he discusses (even if those systems have very few required parts) because the parts of those systems are precisely tailored to each other. Indeed, Behe builds such tailoring into his definition of "irreducible complexity." To assert that a system is irreducibly complex is to assert, not just that the system is complex and irreducibly so, but also that its parts are "well matched" or "precisely tailored" or "highly specified." What this means is that the parts of any irreducibly complex system are able to interact with each other in the way needed for the system to function as well as it does only because they possess a number of very specific properties. Again, the mousetrap is a good example. Its parts are very well matched in the sense that even relatively minor changes in their shape, size, or other characteristics would result in the system's failing to function or failing to function as well. For example, if the holding bar were even slightly shorter, the hammer could not be armed, and hence no mice would be killed. Or if the spring were wound less tightly, perhaps only small or sickly mice would be killed by the hammer's blow.

According to Behe, the problem that such specificity creates for indirect evolution is analogous to the problem one would face if one tried to make a mousetrap from parts that served other purposes:

> In your garage you might have a piece of wood from an old Popsicle stick (for the platform), a spring from an old wind-up clock, a piece of metal (for the hammer) in the form of a crowbar, a darning needle for the holding bar, and a bottle cap that you fancy to use as a catch. But these pieces couldn't form a functioning mousetrap without extensive modification, and while the modification was going on, they would be unable to work as a mousetrap. (p. 66)

What Behe either does not realize or neglects to make explicit, however, is that specificity alone does not suffice to rule out as unlikely a crucial role for the sort of indirect route under consideration. For suppose the precise tailoring of the parts of a system enables it to function better than it otherwise would, but is not required for it to function. Then parts that, because they previously served different purposes, are not precisely tailored to each other could nevertheless begin to function as a system (and be required for that system to function); and then direct evolution could over time improve the system's performance by gradually tailoring the parts until eventually the system becomes irreducibly complex by Behe's definition–it is made up of precisely tailored parts, at least two of which are required for the system to function.

So what Behe needs to rule out the sort of indirect route being considered is a system that cannot function at all unless its parts are well matched. He needs, in other words, not just specificity, but what I will call "irreducible specificity." To claim that the specificity of a system is irreducible is, roughly, to claim that the system would effectively cease to function if its parts were modified. This definition is, however, only a first

approximation, because it treats irreducible specificity as a qualitative concept, whereas in reality it is quantitative–it comes in degrees. No system is such that any modification of its parts, no matter how slight, would result in a total loss of function. For instance, some changes could be made to the parts of a mousetrap that either would not affect how many mice it kills or would make it a less effective but not completely ineffective trap. (For example, the platform could be larger or the spring could be wound slightly looser.) Yet we would still want to say that the mousetrap exhibits a high degree of irreducible specificity because so many relatively small changes in the size, shape, or even composition of its parts would result in a total loss of function. Thus, the crucial issue is whether the parts of Behe's biochemical systems exhibit a sufficiently high degree of irreducible specificity to rule out as unlikely indirect routes in which parts that evolve separately for different purposes unite to form a system with a new function. I suspect they do, but I lack the scientific expertise to make confident judgments about such matters.

Unfortunately for Behe, however, not even (a high degree of) *irreducible* specificity can rule out as unlikely a different sort of indirect Darwinian route. The sort of route I have in mind occurs when an irreducibly complex and irreducibly specific system S that serves function F evolves from a precursor S* that shares many of S's parts but serves a different function F*. Notice that parts that S and S* share and that are required for S to perform F need not be required for S* to perform F* even if they contribute to F*, and parts that are irreducibly specific relative to F may be only reducibly specific relative to F*. Thus, both S* and the specificity of its parts may have been gradually produced via a direct evolutionary path. Then one or more additional parts are added to S*, resulting in a change of function from F* to F. And relative to F, the parts and their specificity, which had not been essential for F*, are now essential.

Once again, it is possible that Behe has good reasons for believing that this sort of indirect route is very improbable; but if he does, he has not told us what they are. For we have already seen that great complexity (in the sense of a multiplicity of parts) will not help rule out a crucial role for indirect routes like this if most of that complexity is reducible, for then a direct route can add that complexity to a minimal system that was produced indirectly. But even if a system is irreducibly very complex, it is not clear why that would make this sort of indirect route unlikely. Behe does offer additional reasons for rejecting indirect Darwinian paths that are based on the individual features of the systems he discusses. For example, he argues that certain parts of the cellular transport system would actually be harmful to the cell if they were isolated from the other parts of the system (p. 112). But these reasons do not by themselves eliminate the sort of indirect path under consideration. So Behe clearly owes us an argument. (Additional arguments are needed to rule out indirect routes involving changes in mechanism.)

C. *Are Direct Routes Impossible?*

Recall that Behe holds that, whether a system has only two parts or a hundred, its production by a direct Darwinian route is impossible if it is irre-

ducibly complex. His argument is quite simple. Suppose that parts A and B are each required for some system to perform its function F. Then without some change in function or mechanism, it is impossible for Darwinian evolution to produce that system by first producing A and then later adding B (or vice versa). For that would require that A be preserved until B appears. But without B, A will not perform some function other than F since we are assuming no change in function and A will not perform F because B is required for F to be performed. Thus, A will perform no function at all, which means A will not be selected and so will not be around when B appears.

I believe that this is a bad argument for several reasons. To distinguish these reasons, it is helpful to distinguish two different types of direct paths. The first, which I will call a *simple* direct path, occurs when there is no change in function or mechanism (making the path direct) and, while parts are added along the way, no parts are subtracted (making the path simple). The second, which I will call a *complicated* direct path, occurs when there is no change in function or mechanism, but parts are both added and subtracted along the way. There is a sense, of course, in which such a path is "more circuitous" or "less direct" than a simple direct path. But keep in mind that, for Behe, directness depends only on there being no change in function or mechanism. So what I am calling a complicated direct path is indeed direct by his definition. (Notice we could also distinguish simple and complicated *indirect* paths if we were so inclined.)

The *possibility* of an irreducibly complex system's being produced by a complicated direct path is fairly obvious. For example, an irreducibly complex two-part system AB that performs function F could evolve directly as follows. Originally, Z performs F, though perhaps not very well. (This is possible because, from the fact that AB cannot perform F without A or B, it doesn't follow that Z cannot perform F by itself.) Then A is added to Z because it improves the function, though it is not necessary. B is also added for this reason. One now has a *reducibly* complex system composed of three parts, Z, A and B. Then Z drops out, leaving only A and B. And without Z, both A and B are required for the system to function.[30]

One might object that in this scenario AB is not really produced directly because, although the route from Z to AB involves no change in function, it must involve a change in mechanism since interaction between two distinct parts is essential to AB's mechanism and so Z cannot function by the same mechanism as AB. If this is right, then no complex system of interacting parts, whether irreducibly complex or not, could evolve both gradually and directly from scratch. An indirect route would always be required in order to get started. But that's no problem for this model. For I can simply stipulate that Z itself is composed of two parts (Z1 and Z2), which do perform AB's function by the same mechanism, and then add that Z1 drops out first, followed by Z2. Of course, Behe has not told us anything about how he individuates mechanisms. But surely he cannot require identity of parts in order to have identity of mechanism. (Otherwise his claim that direct routes are impossible would be trivial.) Furthermore, there is no reason why an indirect Darwinian route could not produce Z, since it has only two parts and, we may suppose, is not irreducibly specific. For example,

perhaps Z1 originally performed F by itself, but by a different mechanism. Then a mutation produced Z2, which united with Z1 to better perform F, but by a new mechanism–the same mechanism used by AB.

Another possible objection to this model of a complicated direct route is that it ignores the irreducible specificity of the parts of the systems Behe discusses. But adding that AB is irreducibly specific will not make its complicated direct evolution impossible. For from the fact that AB is irreducibly specific it doesn't follow that ZAB is. Maybe ZAB was able to perform F before A and B were precisely tailored to each other. Then A and B gradually became more specific because that improved the function of ZAB. And finally, with the loss of Z, that specificity becomes essential for AB to function.

One final objection to this model is that it explains why parts are added, but it does not explain why they drop out. One cannot just appeal to random variation here, because the issue is not why Z might disappear in one organism, but why its absence would become normal for a whole population of organisms. This too, however, is easy to explain. Just as adding a part can be advantageous to an organism and for that reason make that organism more likely to reproduce, so can eliminating a part. For example, if AB functions better than ZAB or requires less energy to produce or maintain, then organisms with the simpler system may very well have an advantage over organisms with the more complicated system. Thus, natural selection can result in a biological system's becoming, not just more complex, but also less complex (and in some cases simultaneously more "machine-like"). In this way, irreducibly complex and specific systems could evolve directly from systems that are reducibly complex and specific.

So a complicated direct Darwinian route to irreducible complexity is clearly possible. But what is perhaps not so obvious is that even simple direct routes are possible, at least on the most natural interpretation of Behe's definition of "irreducible complexity." According to that definition, a system is irreducibly complex if its parts (or at least two of them) are required for it to function. The most natural interpretation of what he means by this is that, if any of those required parts were removed from the *contemporary* system–i.e. if they were removed *right now*–the system would cease to function. This is the natural interpretation because he defends his claims of irreducible complexity by an appeal to what would happen to contemporary organisms if parts were removed from their biochemical systems. But if this is the correct interpretation, then Behe is mistaken in thinking that it follows necessarily from the assumption that a system is irreducibly complex that no precursor could perform the same function. In other words, his claim that "any precursor to an irreducibly complex system that is missing [an essential] part is by definition nonfunctional" (p. 39) is false, and not just because the precursor might have parts not in the contemporary system (i.e. not just because of the possibility of complicated direct routes). One reason for this is that some of the parts of the precursor might have properties that are different from the properties those parts have in the contemporary system, properties that enable the precursor to perform the same function as the contemporary system without some of the parts that are essential in that system. A second reason is that the pre-

cursor might have an environment that is different from the environment of the contemporary system. A part that is required for a system to function in one environment may not be required in another.

Behe's mousetrap example can be used to illustrate these points. This system is irreducibly complex because, as Behe says, it needs all five of its parts in order to kill mice. But it doesn't follow that no precursor to such a trap that was missing a part ever killed mice. For example, a precursor trap lacking a catch might still have worked if the properties of its holding bar (specifically, its shape) differed in the right sort of way from the properties of a contemporary holding bar. In the contemporary trap, the catch is needed both to arm the hammer and to detect the presence of a mouse on the platform. But a bent holding bar can arm a hammer without a catch, and if one is careful to arm it very insecurely, then even the movement of the trap caused by a mouse walking on its platform can cause the hammer to be released.[31] Thus, a precursor trap might still have killed mice, at least occasionally, despite lacking one of the parts that is essential in the contemporary trap. It would not have worked as well as the contemporary trap, but that, of course, is not the issue.

Another possibility is that the environment of the precursor trap was different in some crucial respect. For example, suppose that humans used to feed mice by hand and as a result many mice would boldly approach any human being who was holding food. Then the current mousetrap might have had as a functioning precursor a trap lacking two of its parts. Neither a catch nor a holding bar would have been needed because one could have just held the hammer back and then released it when the cooperative little rodent ran over for a handout.[32] Of course, the immediate "environments" of biochemical systems are other biological systems. Biochemical systems are subsystems of larger biological systems which in turn are subsystems of still larger systems and so on. And given that all of these systems change over time, it is certainly *possible* that a system that cannot work in you or me without some part could have worked in our distant ancestors without that part. Admittedly, I have given no real life examples, but Behe is the one trying to rule out functioning precursors *by definition*.

Behe might try to refute my claim that simple direct routes are possible by claiming that I have misinterpreted his definition of "irreducible complexity." Specifically, he might say that his claim that a part is "required" for function means that the system will not work without it in any (realistic) environment and no matter what properties the other parts of the system might (realistically) have. I have no quarrel with Behe interpreting his definition in this way. But once again, this will not solve the real problem. Instead, it will just shift the problem to Behe's premise that the biochemical systems he discusses really are irreducibly complex. As even the mousetrap example at least suggests, it will now be much more difficult to establish that this premise is true. This is especially so because we do not know for sure what our distant ancestors were like. We cannot examine them in a lab, and there is no fossil record of their biochemical systems. Indeed, they may represent biology's ultimate "black boxes."

Since my objections concerning both complicated and simple direct routes establish only conceptual possibilities rather than probabilities, I

believe Behe's most promising reply to these objections would be to admit the possibility of direct routes to irreducible complexity, but claim that, like indirect routes, they are very unlikely to produce the specific biochemical systems on which Behe builds his case. Of course, it would not be enough for Behe simply to *claim* that direct routes to these systems are very unlikely; he would need to argue for that claim. And it is an open question whether or not a good argument is available.

IV

I have shown in this paper that, while Behe's challenge to Darwinian gradualism is both original and interesting, it is incomplete. The number of possible Darwinian routes to irreducible complexity is greater than Behe realizes. His arguments successfully rule out as very unlikely some routes, but a number of direct, indirect, and mixed routes are not eliminated. I myself doubt that further argument can eliminate all of them, but whether or not I am right about that remains to be seen.

A defender of Behe might respond that the improbability of all Darwinian paths is established by the silence of the scientific literature on the issue of how Behe's systems evolved. Behe himself places great emphasis on this silence, devoting an entire chapter of his book to it. And though some of his critics have suggested that the literature is not as silent as Behe would like us to think, I suspect that Behe is correct that no one has even a good guess, or at least not a very specific one, about how systems like the cilium or anti-body diversity might have evolved. Should we conclude, then, that this silence provides substantial indirect evidence for Behe's position that Darwinism cannot explain the development of these systems? In other words, is it likely that the literature would be much noisier if Behe's position were false?

I am inclined to give negative answers to these questions for two reasons. First, the discipline of biochemistry is very young. Its focus has been and for the most part still is on figuring out how biochemical systems work. If, after a great deal more effort is put into discovering the origin of these systems, Darwinians still make no progress, then I will be more impressed.[33] Second, any indirect evidence that the silence of the literature provides in support of Behe's position is offset by indirect evidence against Behe's position. For in assessing the probability that some system evolved gradually, one cannot just examine the specifics of that system. One must also consider what we know about how other systems evolved. And Behe admits, as well he should, that much evolutionary change is both gradual and driven by Darwinian mechanisms. These Darwinian success stories raise the probability that Darwinian evolution produced Behe's biochemical systems, even if we cannot yet specify step by step exactly how that happened. They may not raise this probability much, but surely they raise it at least as much as the silence of the literature on biochemical evolution lowers it.[34]

Florida International University

NOTES

1. *Darwin's Black Box: The Biochemical Challenge to Evolution* (New York: The Free Press, 1996). Page numbers in parentheses in the text refer to this book.

2. For example, Behe's book has been discussed in *Newsweek*, *U.S. News and World Report*, the *New York Times*, and the *National Review*.

3. Behe's position has been defended publicly by William Lane Craig, William Dembski and William Hasker, and Behe has been a featured speaker at several philosophy conferences.

4. *The Origin of Species* (New York: The New American Library, 1958), p. 171; quoted by Behe on p. 39; my emphasis.

5. *River Out of Eden* (New York: Basic Books, 1995), p. 83; the relevant passage is quoted by Behe on p. 40.

6. Behe compares the situation of scientists searching for an explanation of the development of life to a room in which an elephant (intelligent design) stands next to a flattened body while detectives desperately search for clues to the identity of the perpetrator (pp. 192-193). In fairness to Behe, however, I should also note that he *defends* his inference to intelligent design against a number of objections.

7. "Redundant Complexity: A Critical Analysis of Intelligent Design in Biochemistry," *Philosophy of Science* 66 (June, 1999), 268-282.

8. Behe makes this point in the last section of "Self-Organization and Irreducibly Complex Systems: A Reply to Shanks and Joplin," *Philosophy of Science* 67 (March, 2000), 155-162.

9. *Finding Darwin's God* (New York: Cliff Street Books, 1999), pp. 138-139; my emphasis.

10. Ibid., p. 139.

11. Behe says that *almost* all large biological systems are composed of discrete subsystems (p. 38) and for that reason will not count as irreducibly complex by his definition. But the mammalian ear, or more specifically the three bones whose specific function it is to transmit sound waves from the tympanum to the oval window, may be an exception.

12. Miller also offers a number of other examples, including two molecular ones that he believes provide "a head-on refutation of Behe's assertion that evolution cannot account for the development of complex biochemical systems" (*Finding Darwin's God*, p. 150). The first (p. 149) is the cytochrome c oxidase protein pump, which has six parts, and the second (pp. 150-152) is the Krebs cycle, a metabolic pathway that, according to Miller, has twelve required parts. The problem with these two examples is that neither system has a very large number of parts. Of course, what exactly counts as a *single* functioning part of these systems is somewhat problematic. But, as described by Miller, these systems are relatively simple; so their indirect evolution does not undermine Behe's claim that *very* complex systems are extremely unlikely to evolve gradually via an indirect route.

13. <http://www.infidels.org/library/modern/peter_atkins/behe.html>.

14. *Finding Darwin's God*, p. 134.

15. Ibid., p. 135.

16. Ibid., p. 136.

17. *Natural Theology: Or, Evidences of the Existence and Attributes of the Deity, Collected From the Appearances of Nature* (Houston, Texas: St. Thomas Press, 1972), ch. 1.

18. Elliott Sober, *Philosophy of Biology* (Boulder, Colorado: Westview Press, 1993), pp. 34-35. Richard Dawkins holds a similar view; see *The Blind*

Watchmaker (New York: W. W. Norton & Co., 1985), p. 5.

19. Behe's defense of premise (3) also includes a critique of two non-gradualist alternatives to intelligent design, complexity theory and symbiosis (pp. 187-192).

20. H. Allen Orr, "Darwin v. Intelligent Design (Again)," *Boston Review* 21.6 (Dec/Jan 97) <http://www-polisci.mit.edu/bostonreview/BR21.6/orr.html>.

21. The idea of molecules as distinct from atoms was first suggested by Amedeo Avogadro in 1811, nine years after the publication of Paley's *Natural Theology*.

22. Actually, even Miller strays from Behe's definition on occasion. He tends to vacillate between claiming that irreducible complexity can and often does evolve gradually and claiming that, because certain complex biochemical systems do evolve gradually, they are not irreducibly complex. An example of the latter sort of claim occurs when he asserts that "the notion of irreducible complexity is nonsense," on the grounds that "Evolution assembles complex biochemical machines . . . from smaller working assemblies that are adapted to fit novel functions" (*Finding Darwin's God*, p. 150).

23. *Finding Darwin's God*, pp. 140-143.

24. "A Biochemist's Response to 'The Biochemical Challenge to Evolution,'" <http://www.cbs.dtu.dk/dave/Behe_text.html>. This is a longer and updated version of a review of Behe's book published in *Bios* 70, 40-45.

25. Behe also devotes a chapter of his book to the synthesis of AMP, but since he claims that this system is not irreducibly complex, his argument that it could not evolve gradually is distinct from the main argument of the book, and so is beyond the scope of this paper.

26. The immune system is made up of four subsystems, self-toleration being the fourth. Although all four are required for the immune system as a whole to function, Behe never claims that the immune system as a whole is irreducibly complex, presumably because it is not a discrete system (pp. 138-139).

27. "Philosophical Objections to Intelligent Design: Response to Critics," <http://discovery.org/ embeddedRecentArticles.php3?id=445>.

28. I have been equating "very complex" with "has very many parts" because, in the first part of his book, Behe says, "The systems I discuss are complex because they contain *many* components" (p. 48; my emphasis). But perhaps this interpretation is incorrect. Perhaps, when Behe claims in the first part of his book that, as the complexity of an interacting system increases, the probability of an indirect route decreases, he intends degree of complexity to depend, not just on the number of parts in the system, but also on how specific or well matched those parts are. The advantage of this interpretation is that the apparent shift in emphasis from multiplicity of parts in the first part of Behe's book to specificity of parts in the second turns out not to be real. The advantage of my interpretation is that it emphasizes the distinction between these two different ways in which a system can be "complicated," which is important because they suggest different arguments for the position that indirect routes are improbable.

29. This possibility was pointed out to me by Carl Craver.

30. The main point of two articles on Behe's book by H. Allen Orr (an evolutionary biologist) is, I believe, that complicated direct paths to irreducible complexity are possible. Though Orr's description of a direct path to irreducible complexity in his first article ("Darwin v. Intelligent Design (Again)") could be interpreted in more than one way, he clearly suggests a complicated direct path in the second article when, referring to Behe's mousetrap example, he says, "under the Darwinian scenario, all the parts can change through time

and there's no reason to think we started with anything like a holding bar, spring, or platform" ("H. Allen Orr Responds," *Boston Review* 22.1 [Feb/Mar 97]; <http://www-polisci.mit.edu/ bostonreview/br22.1/orr.html>). Orr's point is echoed by numerous critics who make an analogy to the building of a stone arch, which requires scaffolding that is removed once the arch is complete.

31. These claims are easy to demonstrate. Simply remove the catch from a trap and bend the holding bar so that (roughly) the middle of it can be placed just barely under the tip of the slightly curved end of the spring that extends under the hammer when it is armed. This will make it possible to arm the hammer without a catch, and the closer to the tip of the spring the holding bar is placed, the more insecurely the hammer will be armed. Additionally, one can make the trap even more sensitive to pressure on the platform by allowing the end of the holding bar to extend below the platform when the hammer is armed (so that the platform will not lie flat on the floor when the trap is set). If all this is done correctly, even very light pressure on the platform will cause the hammer to be released.

32. One might object that the ancestral trap has just as many parts as its descendent, because a human hand performs the role of the holding bar and a human visual system takes the place of the catch. But since the original trap required these human components to set the trap, the ancestral trap is still simpler.

33. This point has been made in a paper published on the internet that I cannot now locate. My apologies to the author, whose name I cannot remember.

34. I am grateful to Craig Delancey, Bruce Hauptli, Jeffrey James Jordan, William J. Wainwright, an anonymous *Faith and Philosophy* referee, and especially to Carl Craver and William Hasker for very helpful comments on earlier versions of this paper. I am also grateful to Jeffrey Jay Lowder for referring me to several relevant articles published on the internet.