PCID 3.1.6, November 2004

Irreducible Complexity[1]

By Stephen Griffith

Biochemist Michael Behe created quite a stir with the publication of Darwin's Black Box :The Biochemical Challenge to Evolution.[2] In this book, Behe argues that relatively recent advances in biochemistry reveal the existence of hitherto unknown and unexpected biochemical machines and other complex systems within living organisms which exhibit a property Behe defines as "irreducible complexity". What makes the existence of such allegedly irreducibly complex features of living organisms so interesting and important is that, according to Behe, they (or, more precisely, their existence) simply cannot be explained by means of the various neodarwinian mechanisms available within the context of contemporary evolutionary theory, such as natural selection and chance mutation of genes. To put it another way, if Behe is right, neodarwinism, the currently accepted paradigm of evolutionary theory, will need to be abandoned or at least modified beyond recognition if we wish to provide a truly scientific explanation of these features. This in itself would be enough to agitate much of the scientific community, but an even greater source of agitation is that Behe eventually goes on to argue that the only possible explanation of the existence of irreducibly complex biological features of living organisms must be in terms of "intelligent design". Behe's overall position can thus be seen as involving three important claims:

1. There are some features of and systems within biological organisms which are "irreducibly complex".
2. The existence of irreducibly complex features and systems cannot be adequately explained in terms of neodarwinian mechanisms.
3. The existence of irreducibly complex features of and systems within biological organisms can be explained only in terms of "intelligent design".

It is worth pointing out that these three claims do not in themselves constitute a valid argument in favor of "intelligent design", however defined. To formulate such an argument, we would at least need to add a premise to the effect that, e.g.,

(2a) If irreducibly complex features and systems exist and cannot be explained in terms of neodarwinian mechanisms, then their existence can only be explained in terms of intelligent design,

and this claim, like the first three above, would need to be defended as well. Although Behe and other defenders of "intelligent design" have presented arguments for all four of these propositions, they realize that most devotees of modern science are not likely to take Behe's argument for intelligent design seriously unless they can conclusively demonstrate the truth of (1) and (2). The truth or falsity of (1) and (2), however, can only be ascertained if we have a clear grasp of the concept of "irreducible complexity", and a perusal of the popular, philosophical, and biological literature pertaining to this issue reveals a great deal of confusion concerning this concept. This confusion is especially prevalent among those attempting to defend neodarwinism, which understandably undermines many of the arguments directed against Behe's position.[3] The purpose of this paper is the comparatively modest one of (a) attempting to clarify the concept of irreducible complexity, thereby facilitating the task of those who wish to ascertain the truth or falsity of (1) and (2) above, and (b) examining some of the arguments directed against Behe's position in light of this clarification. ((2a) and (3) are well beyond the scope of this paper and must be dealt with on a separate occasion).

Behe first defines "irreducible complexity" in the following way:

> By *irreducibly complex* I mean a single system composed of
> several well- matched, interacting parts that contribute to the basic function, wherein the removal of any one of the parts causes the
> system to effectively cease functioning.[4]

Given his definition of irreducible complexity, it is easy to see why Behe believes that the existence of biological organs and systems which satisfy this definition present (or would present—we should not beg the question concerning the existence of such entities) a challenge to neodarwinian theory. According to this theory, evolutionary change is to be completely explained in terms of two processes, chance genetic mutation and natural selection, together, of course, with the basic laws of chemistry and physics. The process of evolutionary change is initiated when an existing organism undergoes a chance genetic mutation which happens to benefit the organism, where to “benefit” an organism is understood to mean that it confers a selective advantage to that organism in terms of its ability to survive and reproduce.

Since the nature of biochemistry dictates that any one such change will be very slight, it follows that evolutionary change will be relatively slow and gradual and will require a very large number of changes to effect the sorts of evolutionary change allegedly reflected in the fossil record. Given the fact that life on earth has endured for literally millions if not billions of years, neodarwinians believe that there has been plenty of time for such major changes to have occurred.

So why does irreducible complexity present a challenge? The problem is that, by hypothesis, an irreducibly complex organ or system does not acquire its primary function, and thus confers no selective advantage, until it has been completely "assembled", so there is no way for the neadarwinian mechanisms of chance mutation and natural selection to produce or explain them.

Before pursuing this line of argument any further, in order to give a fair and accurate assessment of Behe's position, it will be necessary to discuss certain refinements to Behe's definition and his argument, most of which are at least implicit in his writings. In the first place, we must explicitly mention what might be referred to as the "strict irreducibility" condition, according to which the organ or system "cannot be simplified and [still] retain the level of function needed for selective advantage".[5] The reason for this condition is simply to underscore the meaning and significance of the sort of "irreducibility" involved in irreducibly complex organs and systems.

This condition also has an important consequence concerning the applicability of the concept of irreducible complexity to actual biological examples, and this consequence might at first seem to undermine Behe's argument. For example, there might very well be hypothetical or real examples of systems which perform more than one function, but do not use all of their "parts" in the performance of any of their functions. Suppose that an organ uses (and must use) all of its parts except part x to perform function A, and uses (and must use) all of its parts except part y to perform function B. Since neither of its functions requires all of its parts, can we still regard it as irreducibly complex? Perhaps we could (or should) redefine the organ in question as two distinct organs, each having all but two parts (x and y) in common, both of which would then satisfy our definition of irreducible complexity more readily. Or consider a situation in which an otherwise irreducibly complex organ or system performs only one primary function but contains parts which are either unnecessary for the performance of that function or redundant. (To consider an artifactual analog, suppose that a manufacturer affixes small metal plates displaying the

company logo on each of its widgets. Would this render the widgets no longer irreducibly complex, on the grounds that the metal plates could be removed with no loss of functionality?) one move here would be to redefine the organ or system in terms of some "core" set of parts, so it would be the core which was irreducibly complex. Or perhaps we could redefine irreducible complexity in the following way:

A system of biological elements is "irreducibly complex" if and only if

(1) these elements are "well-matched"

(2) these elements are "integrated"

(3) these elements collectively perform a "function"

(4) no proper subset of these elements, however arranged, can perform this same function

It must be understood that a "system of biological elements" may or may not be coextensive with a particular biological organ or system as traditionally understood. Thus, for example, it could turn out that the mammalian eye in its entirety is not irreducibly complex, but a subset of its parts constitutes a "system" which satisfies our definition. (Perhaps, for example, pigmentation can be ignored for this purpose). Behe's critics sometimes attempt to make use of the possibility of such unnecessary or redundant "parts" in their attempts to respond to his challenge, but there does not initially seem to be any reason not to apply the concept of irreducible complexity to such "core systems", whether or not they correspond to actual biological organs or systems in their entirety.

Another refinement to Behe's definition which is at least worthy of consideration has to do with just how complex an irreducibly complex entity must be. Consider, for example, a knife consisting simply of a blade (the entirety of which, let us suppose, has been sharpened) and a handle. Neither the blade by itself nor its handle can function as a knife, since the handle by itself cannot be used to cut all of the same things that the knife could cut, and the blade by itself cannot in this case be gripped like a knife without cutting oneself. Does it follow from this that the knife is irreducibly complex? Although there is a sense in which the parts are "integrated" and "well-matched" and collectively "perform"[6] a function, and it is clear that the knife cannot "perform" its function if either part is missing, it is also clear from the context that this is not the sort of thing Behe has in mind.

There are several options open to Behe at this point. In the first place, he might simply accept the fact that there can be irreducibly complex entities with very few parts. If he does this, however, it might seem that this would present a problem for his overall thesis, since neodarwinian mechanisms might very well produce entities with so few parts just by chance.[7] But Behe can counter this move in at least two ways. In the first place, he can simply point out that, even if this is so, it turns out to be irrelevant, since, at the biochemical level, even the simplest organs and systems have far more than two or three "parts", and thus continue to be inaccessible to neodarwinian mechanisms, assuming that they are indeed irreducibly complex. In the second place, he can point out that independently of this fact, irreducibly complex entities, both artifactual and biological, admit varying degrees of complexity,[8] so that the more complex they are, the more inaccessible they are to neodarwinian explanations. Since one need not go too far up the scale for the inaccessibility to become overwhelming, or so he could argue, his position is secure. In the third place, he might simply stipulate that organs or systems must reach a certain level of "minimal complexity", a level which turns out to be inaccessible to neodarwinian mechanisms, not by definition, but simply as a result of the laws of probability.

On the other hand, Behe could argue that things like our knife, as well as other relatively simple artifacts (and biological entities, if there be such) are not irreducibly complex at all, since they fail to exhibit the proper sort of function. So what is the right sort of function, within this context?

As previously mentioned,[9] the concept of "function" is notoriously problematic in the context of the philosophy of biology, and we need now to attend more closely to what it must mean when it pertains to those biological organs and systems which are of interest to Behe. In ordinary discourse, the term "function" is essentially synonymous with the term "purpose". The term "purpose", however, is closely associated with the idea of a purposive agent, and since no such agent is directly observed in connection with the development of biological organs and systems, biologists prefer the term "function", which does not seem to carry with it any such troublesome association. Thus, although the use of the term "purpose" is often obviously appropriate in cases involving intelligent agents, in deference to theophobic sensibilities, let us speak of "function" in all cases, even when the term "purpose" would be less awkward.

There are at least three different sorts of situations in which we are inclined to use the terms "function" or "purpose" in connection with physical objects or systems. The first type of case might be

called a "utilization" case, which occurs whenever an intelligent agent realizes that some naturally occurring, undesigned object such as a rock or a stick can be utilized as a weapon or tool, and can to that extent be said to "function as" such a thing, (or, less awkwardly, to "serve a purpose"). Although, of course, for any object whatsoever, it is logically possible for that object to have been designed by an intelligent agent for some particular purpose, both the nature and the existence of "found" objects of this sort can ordinarily be adequately explained entirely in terms of natural forces. There is typically no reason whatsoever to believe that the object so used was designed for that or for any other purpose, so objects of this sort are clearly of no interest to us within the present context.

A second sort of case in which things may be said to have a function (or in this case "purpose"), which we might call an "artifactual" case, is one in which they have been deliberately designed and /or produced by an intelligent agent to be used by that or some other agent for a specific purpose. Things of this sort are generally recognized as designed, even in many cases in which we do not know either what the purpose of the object was or who designed it. The vast majority of artifacts are such that it would be extraordinarily unlikely that objects exactly like them could have been produced by natural forces alone, without the intervention of an intelligent designer, and there is seldom any question as to whether some particular thing of this sort has been designed.

The third and most contentious category, of course, which we might call the "natural" case, consists of living organisms themselves or features of such which have clearly not been designed by those or other living biological organisms, but which nevertheless clearly have a function. Behe's biological examples of purported irreducible complexity clearly fall within this category, as do many biological features which are not irreducibly complex. Unlike artifacts designed by intelligent beings, the purposes of which are clearly known at least by those who have designed them, biological features of this sort need not even be recognized by the organisms which possess them as having whatever functions they have in order to perform those functions. However, they are nevertheless the sorts of things which at least appear as if they might have been designed for a specific purpose (and it is important to note that even arch-Darwinians like Richard Dawkins describe biology itself as “the study of complicated things that give the appearance of having been designed for a purpose”).[10]

Now that we have delineated three ostensibly different sorts of cases in which something might be said to have a function, it should be clear that Behe simply is not interested in cases of the first sort mentioned above. It should be equally clear that he wants ultimately to claim that at least some things in the third sort of case, like things in the second, must have been designed by an intelligent agent. Finally, however, it should be clear that not everything in either the second or third category is irreducibly complex. Intelligent agents can easily design things that are not irreducibly complex, and biological organisms contain many functioning organs and systems which are not irreducibly complex, whether or not they were designed, as many religious people believe, by an intelligent agent. The point here is that the class of irreducibly complex entities is not coextensive with the class of entities which clearly have a function, nor even with that of those which have been or which appear as if they may have been designed. With regard to their significance, what sets irreducibly complex organs and systems apart from other biological entities which might also give at least the appearance of having been designed is that the former, according to Behe, simply cannot be explained by means of Darwinian mechanisms, whereas it is at least possible that the latter can, at least in principle. So where do we draw the line between things which are and things which are not irreducibly complex?

Clearly, when Behe is thinking of something as irreducibly complex, he is thinking primarily of "machines",[11] including biological machines, and machines typically do things. Every example he gives, biological or otherwise, involves some sort of activity or process (catching mice, clotting blood when appropriate, transporting molecules through the vascular system, etc.)

Behe's first and favorite example of irreducible complexity, which he apparently regards as paradigmatic, is that of a simple mousetrap. Now mousetraps, unlike Venus flytraps, for example, are neither biological organisms nor features thereof. We might therefore ask why Behe uses a human artifact like a mousetrap as a paradigm case, rather than, for example, his other favorite example, the bacterial flagellum. An uncharitable reviewer might suspect that he is trying already to suggest what he ultimately concludes, i.e., that all irreducibly complex things must have been designed, since even the simplest mousetraps are clearly designed and produced by intelligent human agents. A better reason for this choice, however, which actually tends to support his ultimate thesis, is simply that real biological examples of

purported irreducible complexity are far too complex to be of much use in explaining the concept. Let us therefore consider the mousetrap.

A standard mousetrap of the sort considered by Behe consists of five main parts together with an indefinite number of staples or other fasteners to hold it together. In particular, it includes a platform, a hammer, a catch, a spring, and a holding bar. If any one of these parts is missing, or even significantly damaged, the trap simply will not work. Moreover, the parts must be "well-matched" in order for it to function. If the spring is too strong, for example, the trap cannot readily be set, or will break the holding bar when it is, and if it is too weak, it will neither kill nor entrap a mouse. Such a mousetrap thus satisfies Behe's definition of "irreducible complexity".

But what does this have to do with evolutionary biology? To answer this question, we must attend to the analogy which Behe is obviously drawing between mousetraps and the sorts of biological organs and systems which Behe regards as irreducibly complex. But before going into any detail, it might be worth pointing out that for any analogy, there must necessarily be corresponding disanalogies as well,[12] and although these disanalogies need not detract from the point of the analogy itself, there are times when it might be helpful to remove one or more of them. One such disanalogy in Behe's mousetrap example is that we tend to think of mousetraps as individual, independent devices rather than as parts of anything any more complex, whereas the biological organs and systems which Behe regards as analogous to mousetraps (bacterial flagella, blood clotting cascades, vesicular transport systems, etc.) tend to be integral "parts" of even more complex biological organisms. Although this particular disanalogy does not materially affect Behe's argument in itself, the point of this argument might nevertheless be made more clearly if we construct an imaginative scenario involving a mousetrap which is more closely analogous to Behe's biological examples. Let us therefore suppose that someone designs a very complicated mechanical device the purpose of which is to rid one's house of pests. Although it is designed to capture and kill all sorts of pests, its primary targets are insects and mice, so it is dubbed a "Mouse and Insect Killing Entity" (or MIKE, for short). As it turns out, one of MIKE's key features is a mousetrap of just the sort that Behe describes, except that unlike ordinary mousetraps, it is integrated with the rest of MIKE in such a way as to enable MIKE to remove and dispose of dead mice, rebait and reset the trap, and so on. MIKE, being far more complex than a simple mousetrap, is thus more nearly analogous to a living organism than an

ordinary mousetrap is, and the mousetrap itself, understood as an integral functioning part of a more complex functioning entity, is now more nearly analogous to the sorts of biological organs and systems which Behe regards as irreducibly complex. We are thus in a better position to illustrate the relevance of this analogy to evolutionary theory.

Let us try to imagine how a mousetrap, understood now as an integral part of MIKE, might have come into existence by Darwinian or neodarwinian means. The problem which immediately presents itself is that evolution requires natural selection, and natural selection requires functionality. Catching mice, in this instance, is a function which we can suppose would be selected for, since it clearly enhances MIKE's performance in ridding one's house of pests (and would thus be analogous to enhancing the fitness of a biological organism). But none of the parts of the mousetrap, nor any proper subset of these parts, however arranged, would serve this function, by hypothesis. (This is simply part of what it means to say that the mousetrap is irreducibly complex). But then none of these parts individually, nor any proper subset of them, however arranged, would be selected for, and no mousetrap would come into existence in this manner. MIKE's owner, of course, an intelligent designer, might add the parts one at a time with the intention of adding a mousetrap to his device, but neodarwinian evolution does not (and cannot) "look ahead" in this way. Even if a platform, or a catch, or a holding bar, somehow became attached to MIKE just by chance, evolution would tend not to preserve them, since, by hypothesis, they have no function by themselves. Behe's point, of course, is not that mousetraps or more sophisticated devices like MIKE cannot evolve by completely natural means, since we already know that such things do not evolve in this way, but are designed and built by intelligent human agents. His point is that there are things in nature, such as the bacterial flagellum, that are like mousetraps in this respect, (i.e., in the respect that evolutionary mechanisms could not preserve and sequentially arrange sequentially appearing, otherwise functionless parts so as to form a functional organ or system) and thus cannot have evolved in a neodarwinian manner.

As previously mentioned[13], Behe's position cannot be fairly and accurately assessed without mentioning certain refinements to his definition of irreducible complexity and to his argument against the possibility of explaining, by neodarwinian means, the existence of entities which satisfy this description. Refinements to his definition having been discussed, it is now important to point out a subtle distinction in Behe's argument between direct and indirect Darwinian pathways.[14] According to Dembski, a direct

Darwinian pathway is simply "one in which a system evolves by natural selection incrementally enhancing a given function"[15], where the function remains the same. Given our common understanding of Darwinian mechanisms, whereby slight mutations are "selected for" only if they confer some advantage, and given the fact that irreducibly complex organs and systems, by hypothesis, are functional only after they have been completely assembled, it follows that it is conceptually impossible for such organs and systems to evolve by a direct Darwinian pathway. It is important to note that Behe is not arguing that it is logically impossible for such things to come into existence naturally, since almost anything is logically possible.[16] Thus, showing that such things are logically possible cannot count against Behe's point, which is that unless it can be shown that such things can come into existence as a result of a long series of slight, gradual changes, their existence cannot be given a straightforward Darwinian explanation. For this reason, many of Behe's best informed critics try to argue that irreducibly complex organs and systems can evolve by means of indirect Darwinian pathways, in which the organs or systems in question have different functions at different times, each of which is selected for at that time.

Behe's response to this sort of position is twofold. In the first place, he says that, although we are discovering more and more organs and systems which appear to be irreducibly complex, neither the fossil record nor biochemical research has yielded any evidence that anything like this has actually happened, or even can happen. In other words, there is no significant empirical evidence to back up the claim that any irreducibly complex organs or systems have evolved by indirect Darwinian means. The best neodarwinians have been able to do is to cite a few isolated examples of relatively simple organs which superficially resemble structures which constitute integral parts of more complex organs which have a different function, and even in these cases it is far from obvious that these organs have actually served as evolutionary precursors of more highly developed organs having a different function. To a neutral observer, this seems very much like "grasping at straws", especially since there are so few examples, and as Behe points out, "examples of irreducible complexity can be found on virtually every page of a biochemistry textbook."[17]

In the second place, Behe argues that the inherent likelihood that anything like this could occur is too small to be taken seriously. Neodarwinians, of course, are wont to dismiss this claim simply by pointing out that nature has had literally millions if not billions of years to complete tasks of this sort, and that

extremely improbable things are bound to happen if we wait long enough. Behe's defenders, on the other hand, to use one of their favorite examples, can point to the fact that it would be logically possible for something like the Mt. Rushmore monument to be created by the completely random processes of erosion by wind and water, but nature has never produced any such thing, despite the fact that outcroppings have had even more time to erode than life-forms have had to evolve, and Mt. Rushmore is much less complex than most of the allegedly irreducibly complex organs and systems of interest to Behe, such as the bacterial flagellum. Before going any further, however, let us consider some of the actual responses to Behe which have occurred in the literature.

For various reasons (lack of time, lack of interest, etc.), most neodarwinians have not bothered to respond to Behe. Among those that have, many have simply not understood his point.[18] Others seem to have understood his main point, but in attempting to respond to it, have intentionally or unintentionally mischaracterized his concept of irreducible complexity.[19] Although it would be neither appropriate nor necessary to attempt to deal with all of these responses, consideration of some of the more plausible responses might be helpful within the present context.

Scaffolding[20]

One of the common responses to Behe is that, even if we agree with him that there are some biological organs and systems which are irreducibly complex, he fails to consider all the ways that such an entity might come into existence. According to some of these critics, one such way, discussion of which is apparently motivated by Behe's comment that irreducibly complex features of organisms would need to come into existence "all at once" in order to be selectable by natural selection, is by way of "scaffolding".[21] Consider a simple stone arch, of the sort pioneered by the ancient Romans. Such an arch, according to these critics, might be regarded as irreducibly complex, in the sense that its structure might be such that if any one stone is removed, the arch will collapse, and thus be unable to perform its function, whatever that function might be. And yet we know that such arches are constructed gradually, one stone at a time, with the aid of scaffolding. The suggestion, of course, is that a similar process may be at work in the case of irreducibly complex biological systems and organs. The idea is that, while the system in question was "evolving", other parts of the larger organism, which have now ceased to exist (just as we discard

scaffolding when a construction job has been completed) were holding the parts together and thus enabling them, together with the scaffolding, to collectively perform a selectable function.

There are several difficulties with this analogy. In the first place, is it really true that Roman arches are irreducibly complex in Behe's sense of the term? According to Behe's definition, the parts of an irreducibly complex organ or system must be "well-matched" and "integrated", must "contribute to the basic function", and must each be essential to the performance of that function in the sense that the organ or system will cease functioning if any one of its parts is removed. The stones which comprise a Roman arch do appear to be "well-matched", in the sense that they have the appropriate size, shape, weight, hardness, etc. They would also appear to be "integrated" in the sense that they fit together and are situated appropriately to constitute an arch. But is the "function" of an arch the right sort of function? If Roman arches are irreducibly complex, should we not confer a similar status on "natural arches"? After all, multi-parted "natural arches" can also serve a purpose (e.g., an easier way to get from here to there, shelter from a storm, etc.), and might appear to be "irreducibly complex" in whatever way architectural arches are, since they too would collapse and lose their function if one part were removed, and yet no one thinks they were designed. Natural arches are relatively rare, far more so than the biological organs and systems which Behe regards as irreducibly complex, but they do exist, and are typically created when hard rocks which happen to be arranged in a somewhat "arch-like" manner are initially supported by earth or softer rocks which eventually erode away, leaving the natural arch in place.[22] Such an arch thus comes into existence simply as a result of the ordinary forces of nature (gravity, erosion, etc.) acting on a set of randomly placed rocks and the geological media in which they happen to be situated. But would Behe regard them as irreducibly complex? Surely not. For one thing, with regard to the degree of complexity involved, a natural arch is to a Roman arch as the proverbial "Man in the Moon" is to Mt. Rushmore, the latter of which is often cited by intelligent design theorists as something which has clearly been designed and produced by an intelligent agent, and would be recognized as such even if we had no idea who was portrayed or who the sculptor was. But even if the degree of complexity involved in a Roman arch, as opposed to a natural arch, enables us to conclude that the Roman arch, but not the natural arch, was designed, it does not follow that the Roman arch is irreducibly complex. Irreducible complexity is a sufficient condition, not a necessary condition, for design. So is a Roman arch irreducibly complex, or not?

The key here is to see whether Roman arches do or do not have the right sort of function. But to return to our current example, do arches really "do" anything? Yes and no. Perhaps they do something (e.g., physically support things) in a truncated sense of "do", or perhaps an arch counts as a machine in a limiting case of the term. But they are certainly not typical machines. Biology can also provide cases like this. Consider the skeletal systems of vertebrates. Skeletal systems clearly have an important biological function, and it is not "repugnant to the intellect", to borrow a phrase from medieval philosophy, to suppose that they were designed by the Creator. But are they irreducibly complex in Behe's sense? It is possible to argue that they are, since the parts appear to be "well-matched" and "integrated", since they collectively perform a "function" in at least some sense of the term, and would not do so (or do so as well) if one or more bones were removed. On the other hand, it could be argued that this is simply not the right sort of function to satisfy Behe's definition. Although Behe does not spell out what he means by "function" in his definition of irreducible complexity, it seems clear from his examples that he is thinking of machines that respond to certain sorts of situations by engaging in certain sorts of activities or processes that relate to those situations in a way beneficial to the organism in question. Roman arches, whatever else one says about them, do not appear to be machines in this sense of the term, except in a trivial sense.

Whether Roman arches constitute an example of irreducible complexity or not, the "scaffolding" response to Behe suffers from another serious defect. Scaffolding itself is typically a product of intelligent design and often exhibits greater complexity and manifests greater ingenuity on the part of the designer than the thing it is used to construct. Of greatest relevance, of course, is the fact that it may itself be irreducibly complex. But even if it is not, how can the construction and existence of the scaffolding itself be explained by neodarwinian mechanisms? Scaffolding does not typically serve any purpose other than to facilitate the eventual construction of something else which does serve a purpose. But what purpose does the scaffolding serve that could be recognized as such by natural selection? In other words, how does the scaffolding benefit the organism while the organ in question is being assembled? If we say that either the scaffolding by itself or the scaffolding plus the uncompleted arch perform the function ultimately performed by the completed arch itself, why would neodarwinian mechanisms continue to build the arch? It could not be due to improvement in function, since, once again by hypothesis, the arch itself produces no improvement until it is completed. And how do the parts of the scaffolding benefit the organism while the

scaffolding is being constructed? If either the parts of the scaffolding (before the scaffolding is complete) or the completed scaffolding are not benefiting the organism throughout this period of time, why would natural selection preserve them? How, in other words, can neodarwinian mechanisms create scaffolding? Are we to imagine even more complex “meta-scaffolding” being used to construct the scaffolding? Apparently, we are being asked to believe that in the biological analogs to our arch, the organ or system comes into existence as a result of a long sequence of random mutations which just happen to be accidentally coordinated in such a way as to facilitate the creation and assembly of just those "parts" necessary to perform the function performed by the organ or system in question. But why would evolution sequentially select the constituents of the scaffolding, and in the proper order, not to mention the parts of the organ or system itself, and why would it preserve any of these constituents? We are faced with the same problem as before, only worse.

The presence and use of "scaffolding" in the evolutionary pathways leading up to irreducibly complex organs and systems cannot be entirely ruled out on logical grounds alone, for the simple reason that it is in fact logically possible that things have evolved in this manner. As previously mentioned, however, the use of scaffolding in this context seems more difficult to explain using purely neodarwinian means than the existence of irreducibly complex organs and systems themselves. The inherent likelihood is thus quite small, and it is significant that there does not seem to be any empirical evidence (e.g., fossilized "scaffolding") that anything has actually evolved in this manner. Let us therefore consider another sort of response to Behe's position.

Cooptation

Another argument directed against Behe involves the concept of "cooptation" (sometimes called "bricolage" or "patchwork"). The idea here is that the various "parts" of allegedly irreducibly complex organs or systems have been gradually accumulated and preserved because they each served some other function or functions (other, that is, than the function served by the organ or system in question) earlier in the evolutionary development of the organism of which they are a part. To use the mousetrap example, the idea is that each part of the mousetrap comes into existence by chance, happens to benefit the "organism" (i.e., MIKE, in our previous example), and is thus selected for. A part can benefit MIKE either by performing a beneficial function in itself, or by joining with a previously selected and thus beneficial part

and enhancing its function in some way, (although none of these parts, either in themselves or in conjunction with others, can have the function of catching mice, by hypothesis). Once all of the parts have been selected, by purely neodarwinian mechanisms, either individually or as parts of functioning subsets of parts, the parts may come together, once again entirely by chance, to perform a new, beneficial function, (in this case, catching mice). The now completed mousetrap, performing the beneficial function of catching mice, is now selected for, and our story is complete.

To appreciate the point of this response to Behe, we must once again review the problem that the existence of irreducibly complex organs and systems allegedly presents for neodarwinism. According to neadarwinian evolutionary theory, evolutionary change occurs primarily as a result of two naturalistic mechanisms; (1) chance genetic mutation, and (2) natural selection. Genetic changes occur randomly as a result of purely naturalistic phenomena, such as cosmic radiation. Some such changes in an organism are "neutral", in the sense that they neither contribute to nor detract from the ability of the organism to survive and reproduce; some are harmful in one or both of these respects, and some are beneficial, although the percentage that are beneficial is inversely proportional to the relative complexity of the organism, i.e., the more complex the organism is, the less likely it is that there will be a beneficial mutation, so it will subsequently evolve relatively slowly, if at all. When beneficial mutations occur, natural selection makes it more likely that organisms which have these mutations will survive and reproduce (which, of course, is the reason they are referred to as “beneficial”) so the beneficial genetic change becomes more common among organisms of the species in question, and the organism "evolves". The problem is that genetic mutations bring about only slight, gradual changes in an organism, each of which must be selected for individually in order for evolution to occur, but the benefits conferred on an organism by irreducibly complex organs and systems cannot, by hypothesis, occur until all their "parts" have been produced and assembled, and evolutionary processes cannot do this "all at once". The "cooptation" response to Behe is simply that there is no need for evolution to create an irreducibly complex feature all at once, since every gradual step in its creation is selected for for some other reason, i.e., for some other beneficial function it performs other than the function eventually performed by the irreducibly complex organ or system itself.

There are two different ways of interpreting this response. In the first place, we could regard it simply as a rejection of (1) above, i.e., as the claim that there simply are no irreducibly complex features of

biological organisms, since all the evolutionary precursors of such features do in fact have selectable functions. This interpretation, however, simply misconstrues the concept of irreducible complexity, which is defined in terms of the final function of the evolving feature, and if this function cannot be performed by any proper subset of the parts of the feature in question, the cooptation response leaves (1) unscathed. A more plausible interpretation of this response is to regard it as a rejection of (2) above. In other words, we can regard it as the claim that, *pace* (2), evolutionary mechanisms can in fact explain the development of irreducible complexity. In other words, the cooptation response could be that even if there are irreducibly complex organs and systems in nature, their existence does not present any obstacle to neodarwinian explanation so long as each small step in their development is selectable, albeit for some reason other than the role it ultimately plays in the fully evolved biological feature. Here again, as in the case of the "scaffolding" response, it is indeed logically possible that irreducibly complex organs and systems may have evolved in this manner, but here again it seems highly implausible to suppose this[23], and there is no evidence that this has happened, unless we regard the very existence of irreducibly complex features as evidence, which is clearly question-begging. Let us therefore move on to a third response to Behe.

Incremental Indispensability

Biologist Allen Orr asks us to imagine another way in which (he thinks) irreducibly complex organs and systems might come into existence by neodarwinian means.[24] He asks us to imagine that at some stage in its evolutionary development, an organism has a "part" that "does some job", but "not especially well". A chance mutation then adds another part which happens to enhance the performance of the part in question, so that it does its "job" more efficiently. After some time (and many slight, chance, but beneficial mutations), some parts of the original organ change in such a way that the "new" part becomes not just advantageous, but essential to the proper performance of the evolved organ in question. For the sake of argument, let us suppose (implausibly, perhaps) that every part of the evolved organ is not only essential but has also come into existence in this way. The evolved organ will then be irreducibly complex, even though, by hypothesis, it has gradually evolved, albeit indirectly, in a completely neodarwinian manner.

It might appear that Orr has succeeded in explaining the possible existence of irreducibly complex biological features in purely neodarwinian terms, but appearances, as we all know, can be deceiving. Orr

asks us to imagine a scenario in which a functioning organ evolves into an irreducibly complex organ that performs the same function, only more efficiently. If the original organ or part is itself irreducibly complex, this sort of evolution does no violence to Behe's position, since Behe might very well agree that a bacterial flagellum, for example, might evolve in this way into a more "souped up" version, (having greater "bugpower", perhaps). But if the earlier version of the organ, like the "evolved" organ, was also irreducibly complex, Orr has not gained any ground against Behe, since he must now show how the original organ evolved. What Orr needs to show is not how evolution can improve on the function of irreducibly complex organs and systems, but how such things come into existence in the first place. If, on the other hand, the original organ is not irreducibly complex, then Orr must claim that all irreducibly complex organs have non-irreducibly complex evolutionary precursors which perform the same function. (If they perform a different function, then his objection simply reduces to the "cooptation" response previously discussed). The situation here is similar to that involved in the "scaffolding" and "cooptation" responses to Behe. On the one hand, the scenarios envisioned cannot conclusively be ruled out on logical grounds alone. In other words, they are "logically possible" in the broadest sense of the term. On the other hand, they seem highly implausible, and there is no empirical evidence that anything has actually evolved in this way. The discussion here is reminiscent of the common claim that mammalian eyes have evolved from "light-sensitive spots", a claim usually offered with seeming confidence but no scientific research to back it up, (and no admission that "light-sensitive spots" are already very complex, if not irreducibly so)

Reducible Complexity

John MacDonald, as previously mentioned,[25] appears to have initially misconstrued Behe's position concerning irreducibly complex entities and processes, because he thought that by devising functional mousetraps with similar but fewer parts than the standard mousetrap described by Behe he could refute Behe's argument. Behe, however, nowhere states that an irreducibly complex system is one which does a job which cannot be done with a device having fewer parts. What he says is that an irreducibly complex organ or system cannot itself perform the same function if any one of its parts is removed (the other parts remaining unmodified), and the fact that some other, simpler system can do the same or a similar job with fewer parts is irrelevant. In a later response to Behe, MacDonald seems to realize his mistake, and endeavors to show how devices consisting not simply of fewer parts, but of various proper

subsets of the parts of Behe's mousetrap, slightly modified, can serve as functional mousetraps.[26] More specifically, he ingeniously shows us how, by removing one part of Behe's mousetrap at a time, while simultaneously modifying one or more of the remaining parts (to compensate for the loss of the part in question), we can generate a series of progressively simpler mousetraps, all of which are nevertheless functional. This is apparently intended to demonstrate how functionality might be preserved throughout a series of gradual, slight changes, as required by neodarwinian evolutionary mechanisms.

There are several problems here. In the first place, MacDonald's example does not constitute a counter-example to Behe, since Behe's contention that an irreducibly complex entity will not perform its function if any of the parts is removed still stands. It still stands because Behe is clearly talking about removing a part and leaving the remaining parts unchanged. Modifying one or more of the parts is "cheating" in this case, since there is an important sense in which they are not really the same "parts" anymore, and in this respect is tantamount to simply imagining a completely different entity with fewer parts, as he did in his original response. In the second place, the modifications which MacDonald asks us to imagine are obviously designed; they are made precisely for the purpose of compensating for the missing part. Finally, as we shall see below, we cannot assume that a gradual sequence of changes from a more complex entity to a less is the same in every relevant respect as the corresponding sequence of changes from a less complex entity to one which is more complex (more about this below)[27]. Although we can imagine MacDonald's scenario occurring in reverse, doing so requires us to imagine considerable coordination between the modification of an existing part and the addition of a new one, and it is not entirely clear that this can be done in a neodarwinian manner. To return to our previous example, suppose that MIKE initially incorporates a one-parted mousetrap. Are we now to suppose that an additional part, just by chance, appears without any function, and that the original part, just by chance, modifies in such a way as to not only enable the two parts to work together to serve the same function as before, but also to perform it better? Or, alternatively, are we to suppose that the original part modifies itself in anticipation of acquiring the appropriate new part? Once we arrive at a functional two-parted trap, of course, the trap can be selected for by natural selection, but in either case, it seems that there would need to be multiple coordinated changes[28] to ensure that MIKE continues to function throughout the changes. (And what performs the function in question while the changes in question are taking place?) The sort of scenario

proposed by MacDonald, like those previously discussed, cannot be conclusively ruled out on logical grounds alone, but it is not at all obvious that it is suggestive of a satisfactory neodarwinian mechanism for developing irreducibly complex biological features.

Paul Draper

As mentioned previously[29], Paul Draper avoids many of the mistakes committed by Behe's critics, and he also recognizes and acknowledges some of the subtleties in Behe's position, such as the distinction between direct and indirect evolutionary pathways. His general criticism of Behe is that Behe simply has not considered all the logically possible ways that irreducibly complex entities could be created by neodarwinian means. Since most of Draper's scenarios are essentially combinations and permutations of scenarios previously discussed, it will not be necessary to discuss them in detail.

Draper wants to know why a fairly simple irreducibly complex organ or system could not be created by means of an indirect pathway and then refined *ad infinitum* by means of a fairly simple, garden variety direct evolutionary pathway.[30] This, of course, is simply cooptation followed by ordinary evolution, and, as previously mentioned, the problem is that there are far too few examples of indirect pathways, even at a fairly simple level, to account for even a tiny fraction of the examples of irreducible complexity at the biochemical level.

Draper also suggests that an essentially useless genetic mutation could simply be preserved until such time as it became functional (presumably as a result of other changes in the organism).[31] The answer is, once again, that this is indeed logically possible, but evolutionary theory itself counts against it, since natural selection frowns on useless biological features, and, once again, there appear to be no actual examples of this[32].

Draper also suggests that relatively ill-matched parts, which have evolved and been selected for separately, might nevertheless collectively perform some new function.[33] Here again, this is logically possible, but here again, the first stage in the proposed process is highly implausible, at best, and there are few if any incontrovertible actual examples of anything like this happening.

Draper then suggests that we consider the possibility of "a different sort of indirect Darwinian" pathway in which "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*,[34], another version of

cooptation. Draper considers several other possibilities as well, but what they all have in common is that (1) they are logically possible, (2) they are highly improbable, and (3) there are no or very few incontrovertible cases in nature where anything like this could have happened. In view of the aforementioned ubiquity of irreducible complexity at the biochemical level, it would thus appear that Behe has little to fear from Draper's examples.

It seems inevitable that neodarwinians will continue to indulge in imaginative speculation concerning logically possible indirect Darwinian pathways, and it will be impossible for intelligent design theorists to counter each of them individually. As it turns out, however, it may not be necessary to do this. William Dembski has come up with seven "probabilistic hurdles" which any such speculative hypothesis must face, and it is not at all obvious that it is even reasonable to think that these hurdles can be overcome. It is well worth quoting them in their entirety:

(1) Availability: Are the parts needed to evolve an irreducibly complex biochemical system like the bacterial flagellum even available?

(2) Synchronization: Are these parts available at the right time so that they can be incorporated when needed into the evolving structure?

(3) Localization: Even with parts that are available at the right time for inclusion in an evolving system, can the parts break free of the systems in which they are currently integrated and be made available at the "construction site" of the evolving system?

(4) Interfering Cross-Reactions : Given that the right parts can be brought together at the right time in the right place, how can the wrong parts that would otherwise gum up the works be excluded from the "construction site" of the evolving system?

(5) Interface Compatibility: Are the parts that are being recruited for inclusion in an evolving system mutually compatible in the sense of meshing or interfacing tightly so that, once suitably positioned, the parts work together to form a functioning system?

(6) Order of Assembly: Even with all and only the right parts reaching the right place at the right time, and even with full interface compatibility, will they be assembled in the right order to form a functioning system?

(7) Configuration: Even with all the right parts slated to be assembled in the right order, will they be arranged in the right way to form a functioning system?[1]

It is important to keep in mind here that when Dembski uses the term "evolve" and its cognates here, he is referring to evolution by Darwinian means, i.e., by means of chance mutations and natural selection. To the extent that we are depending on chance, the probabilities of each hurdle being overcome come into play, and these need not be merely subjective assessments. On the contrary, using the laws of chemistry and physics, as well as those of mathematical probability theory, the hypothetical probability of any one of these hurdles being overcome in any specific case can actually be calculated, at least in principle. Moreover, since these hurdles must be overcome in temporal order, the probability of overcoming any one of these hurdles being overcome can be calculated under the assumption that previous hurdles have already been dealt with, and the overall probability of overcoming all of the hurdles can then be obtained simply by multiplying the seven probabilities together. These probabilities (or perhaps we should say improbabilities, since they are all relatively small) are not of equal magnitude, but it is fairly easy to see that the overall probability that they can all be overcome is almost vanishingly small, which means that it is not very reasonable to believe that it can be done.

Experimental Evidence?

Since the controversy currently under discussion pertains primarily to recent developments in empirical science, one might think that it could be resolved by means of carefully designed experimentation. So far, this has not been done. As of this writing, (with one possible exception, to be discussed below) no one has conducted any experiments designed specifically to test any of Behe's major claims, partly, of course, because so few scientists have taken the trouble to get clear about exactly what these claims amount to. Some neodarwinians, however, have cited a certain class of experiments known as "knock-out" experiments (which were presumably performed to test hypotheses unrelated to Behe's work) in an effort to refute Behe. It is technologically possible to remove (i.e., 'knock out") the genes responsible for certain biological traits so that an organism will develop without the trait in question. The idea here is that if we consider an organism which has a feature which appears to be irreducibly complex and then

[1] Dembski, William, "Irreducible Complexity Revisited", ISCID Archive, January 2004

"knock out" a gene responsible for one of the "parts" of that feature, that feature either will not develop or will not be able to perform its normal function in the resulting organism, assuming that the feature is in fact irreducibly complex in Behe's sense. If however, the organism is somehow able to survive and reproduce without this feature, perhaps by accomplishing the same goal in some other way, this is supposed to show that Behe is mistaken. The case most often cited is an experiment in which the gene responsible for part of the so-called "blood-clotting cascade", which is one of the biological traits which Behe regards as irreducibly complex, was removed from mice.[35] Although the mice were not entirely normal, as one might expect, they did at least survive for a time with a severely compromised blood-clotting cascade, which supposedly shows that mice of this sort could have been evolutionary precursors of present-day mice. Kenneth Miller, in Finding Darwin's God, cites another case which he believes can serve as an "acid test" of the ability of neodarwinian mechanisms to create irreducibly complex biological systems[36]. According to Miller, the system of so-called "lac" genes in bacteria is just the sort of system which Behe ought to regard as irreducibly complex. He then goes on to claim that when not just one but two "parts" of the system are "knocked out" in the manner referred to above, the organism quickly develops (or "evolves", as Miller insists on saying) a substitute mechanism, by what he regards as purely neodarwinian means. Behe disputes Miller's interpretation of the data in these two cases, and although the debate between them concerning this and other issues still rages on, it would appear that Behe has the upper hand in this instance[37]. A third example of this sort, which might very well have been designed with Behe in mind, involved an experiment in which researchers first deliberately rendered bacteria incapable of producing their flagella in the normal way, and them subjected them to conditions in which motility was necessary for them to secure food. The bacteria in this experiment quickly "evolved" a new way of producing flagella, thereby enabling them to survive. Neodarwinians will no doubt cite this experiment as proof of their theory, since no Gepetto was observed, but there are at least two problems here. In the first place, as pointed out in chapter five, the absence of a Gepetto does not necessarily indicate the presence of chance or the absence of a designer. In the second place, as Behe himself has since pointed out, the experimenters in this instance went to great pains to set things up in such a way as to make it relatively simple for the bacteria to "evolve" new flagella.

It does not appear, however, that experiments of this sort, however they turn out, would be the best way of testing the ability of neodarwinian mechanisms to generate irreducible complexity. According to Behe, for example, the genetically altered mice referred to above are barely viable and unable to reproduce, thus rendering them poor candidates for evolutionary precursors.[38] Far from counting against Behe, this seems to indicate not only that the blood-clotting cascade in mice is indeed irreducibly complex, but also that it is necessary for survival. More importantly, even if the genetically altered mice were both viable and fertile, this experiment does nothing whatsoever to show either that or how neodarwinian mechanisms produced or even can produce the current blood-clotting cascade. With regard to the "lac" experiment, it is not entirely clear that the system in question is truly irreducibly complex (Miller alleges only that it is complex), and, as I have argued elsewhere[39], it is not at all clear why Miller thinks that the substitute mechanisms "evolve" by purely neodarwinian means, or how he rules out the possibility that they were designed. What experiments of this type can (or could) show is whether organisms (like MIKE) which normally benefit from an irreducibly complex organ can compensate for the loss of function which occurs when this organ loses a part, and can thus no longer perform its function. But we already know that organisms (like MIKE) can sometimes repair malfunctioning parts, and sometimes have a certain amount of redundancy which enables them to function even when certain parts are not. The real question here, however, is whether a new, irreducibly complex organ can be generated by purely neodarwinian mechanisms, which is a very different question. At one point[40], Behe makes a suggestion which would come much closer to a true "acid test". Why not select some bacterium, for example, which does not have a flagellum, but which (say) closely resembles bacteria that do, cultivate large colonies of these bacteria under tightly controlled laboratory conditions, subject them to "selection pressure" which makes mobility a highly selectable trait, and see if these bacteria "evolve" flagella. Given the sort of power usually afforded neodarwinian mechanisms and the rapidity with which bacteria typically reproduce, there should be plenty of time to submit these mechanisms to a true acid test. If no flagella are produced, of course, this will not mean that evolutionary mechanisms have failed, since we already know that organisms cannot always adapt to their environment at all, and even if they do not do so, we could always claim that they might have if they had had enough time. But if flagella do evolve, we might then be able to say that the mechanisms

proposed by evolutionary theory are in fact consistent with what actually happens in nature (although, even in this case, this would not rule out design).

There are at least two more responses to Behe's position which must be considered before any tentative conclusions can be drawn, not because they are philosophically compelling, but because they are so common. The first of these responses is simply to ignore Behe. This sort of response might be rooted in a sort of "blind faith" that neoDarwinism must be true, so that nothing Behe says can possibly count against it. It could, of course, be argued that such a "faith" is by no means "blind", since the vast majority of the modern scientific community have embraced neoDarwinism, and since there is such a vast amount of physical evidence consistent with it. But this much could at one time have been said of Ptolemaic astronomy, and we know how that turned out. Scientific theories need not, and should not, be rejected whenever they encounter challenges, but both the scientific and the philosophical search for truth require that we be open to the consideration of such challenges when they arise, as those whose responses to Behe have been considered in this paper have been.

The second of these responses is that, barely five decades after the discovery of DNA, molecular biology is still in its infancy, and we should not expect its practitioners to have all the answers yet. In other words, the response is that even if irreducible complexity cannot be explained now, we will surely be able to explain it at some future date, and our current inability to do so (if we even recognize the existence of such an inability) should not be taken too seriously as an objection to currently accepted evolutionary theory. Anyone who seriously claims that he can now give a definitive account of how life began is either a liar or a fool, and anyone who now makes a similar claim about the details of evolutionary development, especially in the early years of life on earth, is lurking in the same vicinity. It must nevertheless be acknowledged, and Behe would be among the first to do so, that modern science represents humankind's most successful effort to understand and explain the physical universe, and if we ever come to understand the origin and development of life on Earth, it will be scientists, not philosophers or theologians, who enable us to do so. The fact remains, however, that if there are irreducibly complex biological organs and systems, as there seem to be, and as is in effect acknowledged by those opponents of Behe's position whose arguments are discussed in this paper, then it is conceptually impossible for them to evolve by means of direct Darwinian pathways. (Otherwise, there would be no need for these opponents to develop such

fanciful scenarios involving indirect Darwinian pathways). Since one of the most salient features of conceptually impossible situations is that they simply cannot occur, no future scientific discoveries involving direct evolutionary pathways could possibly resolve the situation in Darwin's favor. The only empirical data that could resolve the situation would be the discovery that, contrary to all current appearances, irreducible complexity does not actually exist in nature or can be explained in terms of indirect evolutionary pathways that are significantly less improbable than those currently envisioned. At the present time, it seems highly unlikely that any such discoveries are forthcoming, although it is admittedly not currently possible to calculate just how improbable they would be. Under these circumstances, it would seem that the burden of proof should lie with those who must rely on as yet undiscovered and highly improbable situations and processes to make their case.

Where does this leave us with respect to the goals set out at the beginning of this discussion? In the first place, we have seen that it is important not to confuse irreducible complexity with either complexity *simpliciter* or the appearance of design. Things can be extremely complex without being irreducibly so, and things can be designed or at least have the appearance of having been designed without being irreducibly complex. As previously mentioned,[41] Behe is thinking primarily of "machines", which not only do something, but do it "by themselves", in a sense. Once a mousetrap is baited and set, it performs its function, if it performs it at all, without any further input from the purposive agent who set it. This should be contrasted with cases involving a sort of trap which must be continuously monitored and requires someone to pull a string or push a button at the appropriate time in order to catch a mouse. Similarly, mammalian blood clots "by itself", i.e., there is no meaningful sense in which a wounded mammal must itself consciously initiate the process. In the case of bacterial flagella, it could be argued that the bacteria must in some sense initiate flagellar motion, but this involves endowing bacteria with more volitional control than most scientists would be willing to grant them. An artifactual analog would be the sort of control a human being has when driving an automobile; the driver depresses the accelerator, or turns the steering wheel, but does not directly move pistons up and down or manipulate the orientation of the suspension system. (Indeed, many drivers are relatively ignorant of the nature of these processes) All of these things must be kept in mind when we attempt to determine whether something is or is not irreducibly complex.

Once we have become clear as to how irreducible complexity is to be defined, it becomes an empirical matter whether there are or are not examples of irreducible complexity in nature. Within this context, it is important not to be distracted by nitpicking details, such as the previously mentioned fact that there might very well be hypothetical or real examples of systems which perform more than one function, but do not use all of their "parts" in the performance of any of their functions.

Let us suppose that, as it turns out, there are in fact biological systems that satisfy our definition. What then? According to Behe, neodarwinian mechanisms cannot provide a satisfactory explanation of their existence. How can such a claim be adjudicated? In this case, the situation is somewhat problematic. Most of the arguments produced in response to Behe's position either misunderstand (or at least misconstrue) his position, or offer hypothetical scenarios which are logically possible but highly implausible and have no incontrovertible empirical evidence which can be offered in support of them. Ultimately, of course, just as it is a matter of empirical fact whether irreducible complexity occurs in nature, it is a matter of empirical fact whether such systems, assuming that they do occur, have evolved in accordance with neodarwinian mechanisms.

As previously mentioned, so-called "knock-out" experiments are ill-suited for determining the answer to this question, although experiments of the sort suggested by Behe [42] might show that the development of such systems is at least consistent with contemporary evolutionary theory. There are, however, two final points which must be made. The first is that Behe has issued an important challenge to neodarwinism, one which deserves to be taken seriously, and judging from the responses discussed in this paper, even those who do take his position seriously have not yet succeeded in meeting this challenge. Until someone does, the burden of proof must lie with those who support neodarwinism. The second point is that, even if experimentation shows that the development of irreducibly complex systems is logically consistent with neodarwinian principles, this will not imply that this development can be explained in terms of those principles, and it does not rule out the possibility of intelligent design. But that's another story.

---

[1] I am deeply indebted to Michael Behe for reading and commenting on an earlier version of this paper.

[2] Michael Behe, *Darwin's Black Box: The Biochemical Challenge To Evolution,* (New York, NY: The Free Press, 1996)

[3] For a noteworthy exception to this tendency, see Paul Draper, "Irreducible Complexity and Darwinian Gradualism: A Reply to Michael J. Behe", *Faith and Philosophy,* Vol.19, No. 1, January 2002

[4] Behe, *op.cit*., p.39

[5]Cf. William Dembski, "Evolution's Logic of Credulity: An Unfettered Response to Allen Orr", http:www.designinference.com/documents/2002.12.Unfettered_Resp_to_Orr.htm, p.3. Dembski refers to this condition as the "minimal complexity" condition, but I prefer to reserve this terminology for another possible refinement of Behe's definition (see below, p.7).

[6] It is significant that a more perspicacious locution in this instance would be "can collectively be used to perform a function", since knives do not typically perform their intended functions by themselves, but only when directed and controlled by an intelligent agent, or perhaps by a complex machine.

[7] Cf. Draper, ibid., p. 6.

[8] Cf. Behe, I offer the following tentative "evolutionary" definition of irreducible complexity: An irreducibly complex evolutionary pathway is one that contains one or more unselected steps (that is, one or more necessary-but-unselected mutations). The degree of irreducible complexity is the number of unselected steps in the pathway.", "In Defense of the Irreducibility of the Blood Clotting Cascade: Response to Russell Doolitle, Ken Miller and Keith Robison", *Discovery Institute,* July 31, 2000; p.9.

[9] cf. p.3

[10] *The Blind Watchmaker* (New York, NY: W.W.Noton and Co., 1996), p.1

[11] cf., e.g., *Black Box*, pp.4ff.

[12] Since the purportedly analogous objects or situations would otherwise be indistinguishable.

[13] Cf. p.4.

[14] Cf. Draper, *op. cit.,* p.3; Dembski, *op. cit.,* pp2ff.

[15] Dembski, *op. cit.* p.3.

[16] It is for this reason that I prefer to say, as Draper does, (*op. cit.,* p.3), that it is "conceptually impossible" for irreducibly complex things to evolve by means of a direct Darwinian pathway, rather than to say, as Dembski does (*op. cit.,* p.3), that this is a "logical" point.

[17] "Molecular Machines: Experimental Support for the Design Inference", *Cosmic Pursuit,* March 1, 1998, p.6.

[18] Cf. John MacDonald, who seems, initially at least, to have believed that Behe's point was that, in the case of irreducibly complex organs or systems, there could be no simpler organ or system which could possibly perform the same or a similar function. Behe responds to this mistake in "A Mousetrap Defended: Response to Critics", *http//:www.discovery.org/embeddedRecentArticles.php3?id=446.* McDonald has since updated his examples in response to Behe's response at *http//:udel.edu/ ~mcdonald/mousetrap.html* . (See below for commentary concerning McDonald's revised mousetraps examples, pp.22-24.

[19] Cf., e.g., Thornhill and Ussery, who seem to believe that Behe defines an irreducibly complex organ or system as one which has some one part which is essential to its function. ("A Classification of Possible Routes of Darwinian Evolution", *Journal of Theoretical Biology* 203 (2000): 111-116.

[20]This section and the three that follow it borrow a great deal both in organization and content from William Dembski's excellent discussion of this topic in his book *No Free Lunch,* chapter 5, especially pp. 246-267.

[21] Cf., e.g., Thornhill and Ussery, *op.cit.* , and Thomas D. Schneider, "Evolution of Biological Information", *Nucleic Acids Research* 28 (14) (2000): 2794

[22] Other scenarios involving earthquakes or landslides are also possible, of course.

[23]Even neodarwinian H. Allen Orr, a strong opponent of the intelligent design theory, concurs, as he says "You may as well hope that half of your car's transmission will suddenly help out in the airbag department", "Darwin v. Intelligent Design (Again)", *Boston Review,* (December/January 1996-1997:29

[24] ibid., p. 29.

[25] cf. fn. 13

[26] op.cit.

[27] cf. pp. 30-31

[28] It is worth pointing out that the biological analog of "bending a wire" would almost certainly involve a fairly large number of modifications at the biochemical level. It is therefore quite misleading to think that an organ or system would function differently, much less more efficiently, as a result of one simple change.

[29] Cf. fn. 2.

[30] Ibid., p. 11.

[31] Ibid., p. 12.

[32] So-called "junk DNA" is not an example of this, but it would be beyond the scope ofthis paper to explain why.

[33] Ibid., p. 13

[34] Ibid, p.14.

[35] Cf. Russell F. Doolittle, "A Delicate Balance", *Boston Review,* (February/March 1997): 28-29.

[36] Cf. Miller, *Finding Darwin's God*, (New York, NY: HarperCollins, 1999), pp. 145-147.

[37] Cf. Behe, ""A True Acid Test": Response to Ken Miller", *Discovery Institute*, July 31, 2000, and "Comments on Ken Miller"s Reply to My Essays", *Discovery Institute,* January 8, 2001.

[38] Cf. Behe, "In Defense of the Irreducibility of the Blood Clotting Cascade: Response to Russell Doolittle, Ken Miller, and Keith Robison", *Discovery Institute,* July 31, 2000.

[39] "Empiricism and Design", unpublished

[40] Cf. Behe, "Philosophical Objections to Intelligent Design", *Discovery Institute,* July 31, 2000.

[41] p. 8ff

[42] Cf. fn. 31