‘Evidence’ For Evolution Versus ‘Evidence’ For Intelligent Design: Parallel Confusions

Kirk Fitzhugh

Research & Collections Branch, Natural History Museum of Los Angeles County,

900 Exposition Boulevard,

Los Angeles, CA 90007, USA

e-mail: kfitzhug@nhm.org

Correspondence: Kirk Fitzhugh, Research & Collections Branch, Natural History Museum of Los Angeles County, 900 Exposition Boulevard, Los Angeles, CA 90007, USA. Phone: 213-763-3233; fax: 213-746-2999; e-mail: kfitzhug@nhm.org.

Running head: Evidence, evolution, intelligent design

Abstract The popular defense of intelligent design/creationism (ID) theories, as well as theories in evolutionary biology, especially from the perspective that both are worthy of scientific consideration, is that empirical evidence has been presented that supports both. Both schools of thought have had a tendency to rely on the same class of 'evidence,' namely, the observations of organisms that are in need of being explained by those theories. The result is conflation of the 'evidence' that prompts one to infer hypotheses applying ID or evolutionary theories with the 'evidence' that would be required to critically test those theories. 'Evidence' is discussed in the contexts of inferring theories/hypotheses, suggesting what would be possible tests, and actual testing. These three classes of inference being abduction, deduction, and induction, respectively. Identifying the different inferential processes in science allow for showing that the 'evidence' to which theories and hypotheses provide understanding cannot recursively be the 'evidence' supporting those theories and hypotheses. This clarification provides a strong criterion for showing the inability of an ID theory to be of utility in the ongoing process of acquiring causal understanding, that is the hallmark of science.

Keywords Evolutionary biology • Intelligent design • Testability • Abductive inference

Rather than taking on the creationists obliquely and in wholesome fashion by suggesting that what they are doing is 'unscientific' tout court..., *we should confront their claims directly and in piecemeal fashion by asking what evidence and arguments can be marshaled for and against each of them. ...the real question is whether the existing evidence provides stronger arguments for evolutionary theory than for Creationism*. Laudan (1982: 18)

Introduction

When one looks closely at the arguments used to defend what are generally regarded as mutually exclusive fields of study, i.e. evolutionary biology and intelligent design/creationism (ID), there is an interesting parallel. Proponents on both sides tend to garner support for their respective positions by way of the very observations to which evolutionary (e.g. Charlesworth and Charlesworth 2003; Scott 2004; Barnosky and Kraatz 2007; Prothero 2007; but see Fitzhugh 2008a) or ID theories (e.g. Dembski 1998, 1999, 2002a; Meyer 2002a, b; Davis and Kenyon 2004; Behe 2006) are applied. For instance, in his textbook, *Evolution*, Futuyma (2005: 48-49; see also Futuyma 1998) claims that "From the comparative data amassed by systematists, we can identify several patterns that confirm [*sic*] the historical reality of evolution, and which make sense only if evolution has occurred." He then lists eight sources of 'evidence' for evolution: (1) the hierarchical organization of life, (2) homology (= homology-'homogeny' *sensu* Lankester 1870 relations; cf. Fitzhugh 2006a, b, c, 2008b), (3) embryological similarities, (4) vestigial characters, (5) convergence (= 'homoplasy' *sensu* Lankester 1870; cf. Fitzhugh 2006a, b, c, 2008b), (6) suboptimal design, (7) geographic distributions, and (8) intermediate forms. Except for (1) and (5), which are the products of inferences intended to explain particular observations, points (2) - (4) and (6) - (8) refer to those qualities of organisms that one would be compelled to explain by way of the theories under the rubric of evolutionary biology. In a similar vein, ID has been claimed to receive support from our observations in need of being explained by that theory. For instance, Dembski (2002a: 8, emphasis original) states that when "we infer design, we must establish three things: *contingency*, *complexity*, and *specification*." Establishing these three qualities is a direct consequence of the effects to which design is directed.

The claim by biologists and ID advocates alike that observed effects can serve as 'evidence' supporting their alternative views suffers from a fundamental error – that of confusing the 'evidence' one uses as the basis for inferring a theory or hypothesis with the 'evidence' needed to critically assess either. It is a distinction Hanson (1958: 200, note 2, emphasis original; Achinstein 1970) recognized as too often ignored or denied as a consequence of the adoption of hypothetico-deductivism (H-D). There is,

> ...the logical distinction between (1) reasons for accepting an hypothesis H, and (2) reasons for suggesting H in the first place. (1) is pertinent to what makes us say H is true, (2) is pertinent to what makes us say H is plausible. Both are the province of logical inquiry, although H-D theorists discuss only (1) saying that (2) is a matter for psychology or sociology – not logic. This is just an error.... We are discussing the rationale behind the proposal of hypotheses as possible *explicantia*. H-D theorists never raise the problem at all.

A common consequence of not clearly acknowledging the logic of hypothesis and theory formation is that the 'reasons' in (1) and (2) are not recognized as being very different classes of evidence. Evidence used to infer some theory or hypothesis cannot recursively function as evidence supporting the very theories or hypotheses that have been inferred to account for that evidence. The goal of this paper is to provide a brief outline of how to avoid the common mistake of confusing evidence 'suggesting' a particular theory or hypothesis with the evidence for 'accepting' that theory or hypothesis. Establishing this distinction has the benefit of reorienting debate about evolution and ID back to the real issue: applying the accepted tenets of testing as the means of deciding the explanatory merits of theories (Sober 1999). This is not a matter of promoting any kind of demarcation criterion between science and non-science. As was suggested by Laudan (1983; see also Stamos 2007), rather than stressing demarcation, which might be a tentative proposition, we should judge claims on the basis of what can be provided in the way of supporting or refuting empirical evidence. As such, careful consideration is required to make the distinction stated by Hanson (1958), especially when comparing evolutionary biology and ID.

Evidence and Inference

Inference involves a relation between evidence and conclusion – evidence comprises the premises from which one infers a conclusion (Salmon 1984b). To speak of the support for a proposition is to refer back to the premises. For instance, in the classic example of deduction,

[1] All humans are mortal

Socrates is human

Socrates is mortal,

the first two statements form the premises, from which there is only one possible conclusion, separated by a single line, indicating that the argument is deductive. Given that one of the requirements for valid deduction is that true premises must lead to a true conclusion, we have a conclusion that is certain based on the evidence provided. Using the same statements, but in slightly different contexts, we can identify relations between evidence and conclusion in an inductive inference:

[2] Socrates is human

Socrates is mortal

All humans are mortal.

The difference is that the evidence only allows for a probable, not certain conclusion. What might be applicable to Socrates provides no guarantee that all other humans exhibit the same condition. But, in either inference, we can readily identify what stands as evidence supporting, i.e. providing the basis for, the respective conclusions. Deductive conclusions receive unequivocal support, whereas non-deductive conclusions receive tentative support.

The hallmark of science is that its goal is to acquire causal understanding through the critical and continual elucidation of empirical support for theories and hypotheses (e.g. Hempel 1965; Rescher 1970; Popper 1983, 1992; Salmon 1984a; Van Fraassen 1990; Mahner and Bunge 1997; Hausman 1998). The desire to understand unexpected or surprising objects and events is a basic human endeavor. But, the routes to understanding are varied (Ayala 2004). One could claim that various religions, as well as other predilections, e.g. astrology, also function to acquire understanding, albeit by ways that invoke supernatural entities and processes. The standard within the various fields of science that allow for critical evaluations of theories and hypotheses is the process of testing. This does not mean that testing is necessarily a criterion strictly demarcating religion and science in the context of evaluation. For instance, the testing of Biblical hypotheses regarding the age of the Earth or the occurrence of the Noachian flood have been conducted by scientists (Prothero 2007), although such testing is silent on the claims that the events in question were caused by a supernatural entity. One could speak of what would be required to test a theory of the supernatural, framing the subject in the manner by which inferences of tests are performed in any field of science. As indicated in the quote by Laudan (1982; see also Laudan 1983) at the start of this paper, the real issue is to identify what evidence can actually be provided for the support or refutation of a theory or hypothesis, rather than assuming the outright denial of close scrutiny of supernatural claims in the sciences. Of course, to speak of testing means we once again invoke the term 'evidence.' Given that to accept some theory or hypothesis as having the capacity to increase understanding, be it in the realm of evolutionary biology or ID, there must be discernable, empirical consequences that can be articulated. It is these consequences that serve as potential test evidence. What becomes critical is to recognize how 'evidence' plays different roles in the acquisition of understanding. We have to differentiate between evidence that compels us to infer a particular theory or hypothesis from evidence that, in the context of testing, supports or refutes those claims.

While Hanson's (1958) concern that hypothetico-deductivism has tended to ignore how theories and hypotheses are inferred in the first place, this has not prevented considerations of what would be regarded as proper evidence that either supports or refutes a theory or hypothesis as a consequence of testing. A good example of can be found in the following passage by Popper (1992: 132-133, emphasis original), which is worth quoting in its entirety:

> I suggest that it is the aim of science to find *satisfactory explanations* of whatever strikes us as being in need of explanation. By an *explanation* (or a

causal explanation) is meant a set of statements one of which describes the state of affairs to be explained (the *explicandum*) while the others, the explanatory statements, form the 'explanation' in the narrower sense of the word (the *explicans* of the *explicandum*).

We may take it, as a rule, that the explicandum is more or less well known to be true, or assumed to be so known. For there is little point in asking for an explanation of a state of affairs which may turn out to be entirely imaginary.... The *explicans*, on the other hand, which is the object of our search, will as a rule not be known: it will have to be discovered. Thus, scientific explanation, whenever it is a discovery, will be *the explanation of the known by the unknown.*

The *explicans*, in order to be satisfactory (satisfactoriness may be a matter of degree), must satisfy a number of conditions. First it must logically entail the *explicandum*. Secondly, the *explicans* ought to be true although it will not, in general, be known to be true; in any case, it must not be known to be false – not even after the most critical examination. If it is not known to be true (as will usually be the case) there must be *independent* evidence in its favour. In other words, it must be *independently* testable; it will be the more satisfactory the greater the severity of the independent tests it has survived.

I still have to elucidate my use of the expression 'independent', with its opposites, '*ad hoc*', and (in extreme cases) 'circular'.

Let *a* be an *explicandum*, known to be true. Since *a* trivially follows from *a* itself, we could always offer *a* as an explanation of itself. But this would be highly unsatisfactory, even though we knew in this case that the *explicans* was true, and that the *explicandum* followed from it. *Thus we must exclude explanations of this kind because of their circularity*.

Yet the kind of circularity I have in mind here is a matter of degree. Consider the following dialogue: 'Why is the sea so rough today?' – 'Because Neptune is very angry.' – 'By what evidence can you support your statement that Neptune is very angry?' 'Oh, don't you *see* how *very* rough the sea is? And is it not always rough when Neptune is angry?' This explanation is found unsatisfactory because (just as in the case of the fully circular explanation) the only evidence for the *explicans* is the *explicandum* itself. The feeling that this kind of almost circular or *ad hoc* explanation is highly unsatisfactory, and the corresponding requirement that explanations of this kind should be avoided are, I believe, among the main forces in the development of science: dissatisfaction is among the first fruits of the critical or rational approach.

In order that the *explicans* should not be *ad hoc*, it must be rich in content: it must have a variety of testable consequences, and among them, especially, testable consequences which are different from the *explicandum*. It is these different testable consequences which I have in mind when I speak of *independent* tests, or of *independent* evidence.

What we see in Popper's characterization of science as a process for acquiring understanding is that on the one hand we are confronted with surprising effects. Even a tentative explanation of such effects will have the form of conjoining some set of theories and other background information to a hypothesis of initial causal conditions. While the surprising effects serve as the 'evidence suggesting' a hypothesis, any attempt to critically evaluate the hypothesis itself would require the search for evidence wholly independent of those effects, and in fact a class of effects wholly distinct from that in need of being understood.

What is required next is that we identify the different contexts in which evidence plays a role in the process of acquiring understanding. To do this, we must first outline the major types of reasoning employed in this process.

The Acquisition of Rational Understanding

The formal examination of reasoning has traditionally started with descriptions of deduction and induction (e.g. Salmon 1984b; Copi and Cohen 1998). In this case, if a form of inference is not deductive, then it must be inductive, regardless of nuances in form or motivations leading to the inference. Thus, hypothetico-deductivism is seen as characterizing the interplay between induction and deduction. Induction serves to formulate hypotheses and theories, and deduction allows for testing (Gauch 2003; Barnosky and Kraatz 2007; Prothero 2007). Such a characterization is far from accurate. A large part of the problem is that the one term induction does not convey many of the fundamental actions in science, and deduction is often mischaracterized as entailing the process of testing. I have expressed the view elsewhere (Fitzhugh 2005a, b, 2006a, b, 2008a, b) that I subscribe to the perspective developed by Peirce (1931-1935, 1958), where a more accurate picture of how understanding is acquired in the sciences is to recognize at a minimum the interplay between three classes of reasoning, which characterize (1) the developments of hypotheses and theories, (2) formulations of potential tests, and (3) the actual performing of those tests. These classes of reasoning are commonly referred to as abduction, deduction, and induction *sensu stricto*, respectively.

Abductive Inference: Hypothesis and Theory Formation

The inference of hypotheses is probably the most frequent type of reasoning performed by humans. We are constantly confronted with unexpected or surprising observations. In order to begin the process of understanding those observations, we would invoke some tentative explanation, in the form of past causal conditions, that would make the observations seem a matter of routine relative to our background of past experiences. The term that has often been applied to this type of reasoning, especially since the 19th century, is abduction or abductive inference (Peirce 1878, 1931-1935, 1958; Hanson 1958; Achinstein 1970; Fann 1970; Reilly 1970; Curd 1980; Nickles 1980; Thagard 1988; Josephson and Josephson 1994; Hacking 2001; Magnani 2001; Psillos 2002; Godfrey-Smith 2003; Walton 2004; Atocha 2006; Fitzhugh 2005a, b, 2006a, b, 2008a, b, in press), and more recently 'inference to the best explanation' (Harman 1965; Ben-Menahem 1990; Lipton 1993; Douven 2002). We can schematize abduction in the manner used in the previous section for deduction and induction. The relations between evidence and conclusion have the form,

[3]	Theory:	given causal condition *x*, effect *y* will ensue
	Observation (effect):	*y* is the case

Hypothesis: *x* might have been the cause.

Like induction *sensu stricto* (see below), the non-deductive nature of abduction is indicated by the double line separating premises from conclusion. For instance, regardless of the truth of the premises, the conclusion might be false. While the conjunction of some theory with observed effects provides an explanatory account that makes the observations most probable, in the context of that theory, this does not ensure the truth of the hypothesis. As was emphasized by Popper (1992) in the above quote, explanations have the form of accounting for the known by the unknown. In other words, we invoke unobserved past causes, to give us understanding of what we do observe in the present. To say we rely on our past experiences to give us understanding of what we observe in the present is conveyed in **[3]** by way of applying some theory or set of theories, including unstated background knowledge, to observed effects. We rely on the past successes of theories, as explanatory concepts, to give us at least tentative understanding, under the assumption that those theories will be successful in the present.

The schematic form in **[3]** can be applied to a contrived example of abduction in the context of evolutionary biology. One has examined a group of individuals, all of which have the surprising feature of two pairs of legs, in contrast to one pair. To causally account for these observations, the abductive inference of an explanatory hypothesis would have the following form:

[4] Theory: 'descent with modification,' i.e. in the event a novel feature arises and becomes fixed in an ancestral population as a result of natural selection, that feature will be shared among descendant individuals

Effects: these individuals have two pairs of legs

Hypothesis: an additional pair of legs originated as a consequence of some mutation, providing a selective advantage, whereby it became fixed in an ancestral population, subsequent to which the population diversified.

The example illustrates the application of some of the components of evolutionary biology, under the rubric of 'descent with modification,' to the effects to be explained. The inference leads to the tentative hypothesis that a particular set of causal events occurred in the past. Notice, however, that the hypothesis is a vaguely worded construct, given that the theory invoked in this instance is not very detailed.

In contrast to the evolutionary example just presented, one might invoke the alternate theory of 'intelligent design.' Although no specific or formal design theory has been offered by advocates, we can still represent the approach that would be used:

[5] Theory: 'intelligent design,' i.e. a supernatural, intelligent agent has the capacity to invoke new features in living organisms

Effects: these individuals have two pairs of legs

Hypothesis: an additional pair of legs originated in an ancestral population by the action of an intelligent agent, subsequent to which the population diversified.

Notice again the vague nature of this hypothesis. But in both instances, the inferences produce hypotheses that make the observations most probable, given one's acceptance of one of the theories.

Given that the inferences in **[4]** and **[5]** present contradictory answers to the same causal question, would it be possible at this juncture to determine which of the two should be regarded the more plausible? While this point will be addressed more fully later (see **Failed evidence for ID**) in the context of the relevance of independent, empirical support for auxiliary theories to the process of testing, the issue is also relevant to abduction. Invoking either the theory of natural selection or ID incurs accepting auxiliary theories, upon which the former theories are dependent. If any of these theories cannot be supported, which would have been consequences of previous tests, then the plausibility of an abductive inference is compromised. For instance, the theory of natural selection in **[4]** is dependent upon such auxiliary theories as mutation and inheritance, both of which withstood testing. By the same measure, selection has also been subjected to testing (Endler 1986; Mitton 1997; Kingsolver and Pfennig 2007). A critical auxiliary theory in **[5]** is that a supernatural, intelligent agent does exist, from which the theory of ID speaks of a particular type of behavior of that agent. The problem that the ID theory faces is that no independent empirical support exists for the auxiliary theory, indeed the theory has never been tested and appears untestable. Thus, the relative plausibility of inferences in **[4]** and **[5]** would suggest that natural selection provides the more effective explanatory hypothesis, especially given that it would itself be available to being tested.

While the examples in **[4]** and **[5]** pertain to inferences of explanatory hypotheses, abduction is also the inferential form from which theories are derived. Darwin's (1872) theory of natural selection is a classic example of abduction via analogy (Millman and Smith 1997). Darwin was confronted with trying to explain the differential occurrences of shared similarities in natural populations. As there was no available theory, beyond the supernatural, Darwin relied upon what was known of artificial selection. The analogical, and background evidence upon which natural selection was developed can be expressed as follows:

[6]	Background knowledge:	variation / inheritance / differential survival and reproduction
	Tentative theory:	based on what is known of the actions of artificial selection, in conjunction with the above background knowledge, maybe an analogous system of cause and effect relations exists in nature: ***Natural selection*** - *organisms with traits that enhance survival and reproduction will leave offspring with those traits*
	Observations:	there are differentially shared traits among these observed organisms
		--
	Hypothesis:	variation arose in an ancestral population, subsequent to which the traits in question allowed for enhanced survival and reproduction.

The 'background knowledge' of variation, inheritance, differential survival and reproduction, were clearly identified by Darwin (1872: 3-4) as being conditions already known, albeit theories of variation and inheritance were still tentative. By way of analogy, Darwin surmised that the actions of selection as applied by humans to domesticated plants and animals could be extended to nature at large. As is shown in **[6]**, 'artificial' selection serves as the framework for considering a theory of 'natural' selection. It is then a matter of applying this new theory to the observations. The utility of the analogously-derived theory is that it allows for abductively inferring a new class of explanatory hypotheses.

While abduction plays the fundamental role of producing hypotheses and theories, these are constructs that serve very different roles. Theories present statements of cause-effect relations (cf. **[3]** - **[6]**; Fitzhugh 2008). As such, a theory is not a spatio-temporally restricted account regarding a particular cause-effect relation or set of relations. Relative to theories, however, hypotheses are statements claiming that particular causal events occurred in the past that produced a set of observed effects (cf. **[3]** - **[6]**). Contrary to what was claimed by Meyer (1994: 36; see also Meyer and Keas 2003: 139) in relation to the abductive inference of theories, any theory in the field of evolutionary biology could not then be limited to only being "a theory about what in fact happened in the past." Evolutionary theories pertain to past, present, and future phenomena.

Returning to the subject of evidence in the context of abduction, it is the theory (and relevant background knowledge) and effects that provide the basis for the hypothesis provided (cf. **[3]** - **[6]**). In other words, we have the "reasons for suggesting H in the first place." These reasons, i.e. evidence, stand entirely separate from the "reasons for accepting... H" (Hanson 1958). The basis for offering a hypothesis is always contingent on the introduction of some theory that establishes a cause-effect relation that can be applied to the observations in question. As there are no rules of reasoning that dictate what theory is to be used, the inferences in **[4]** and **[5]** are equivalent as far as form is concerned. An important consequence is that inference **[4]** cannot be judged superior to **[5]**, or vice versa, solely on the basis of the 'evidence' that are the observed effects. In the context of the procedures applied in the sciences, however, **[4]** can be deemed superior to **[5]** from the standpoint that of the two theories, only 'descent with modification' has withstood repeated attempts of refutation by way of testing. In other words, the empirical evidence amassed for descent is substantial, whereas no such evidence has yet been provided for supernatural causes. It will be the misinterpretations of 'evidence' as supporting theories and hypotheses that will occupy the latter half of this paper. At this point, we have confirmed Hanson's (1958) and Popper's (1992) views, quoted earlier, that the observations in need of explanation cannot provide a viable basis for defending any theory or hypothesis that explains those observations. Those observations merely provide the basis for starting the quest for understanding. The ramifications for defending evolution and ID are significant.

Deductive Inference: Potential Tests of Hypotheses and Theories

With abduction providing the process of reasoning from observed effects to tentative hypotheses, as well as the basis for introducing new theories, there is next the consideration of critically evaluating the veracity of hypotheses and theories. In other words, we want to be able to determine the effectiveness of a particular theory or hypothesis at providing understanding. It is the process of testing, or at least the potential for testing, that has stood as the principle hallmark of science. While advocates of ID have often held that their point of view is worthy of scientific consideration, this could be entertained insofar as acknowledging what types of test evidence would need to be introduced. In other words, a distinction can be drawn between potential tests that could lead to viable test conditions as opposed to proposed test conditions that could not even be potentially manifested. To say that ID is untestable might imply that no deductive argument can be presented to outline, at a minimum, what would be required to legitimately assess a supernatural theory. This is not to say such a test could ever be accomplished given what is required to test theories. It would appear that this is the limitation referred to by Laudan (1982) in the beginning quote, which leaves any demarcation criterion between science and non-science context dependent. Claims that ID is testable will be addressed below (see **Testability and ID Theory**). Considering the common misunderstanding among evolutionary biologists and ID proponents regarding the nature of 'evidence' confirming or refuting their respective theories and hypotheses, as indicated earlier, we must now examine the formal structure of inferring potential test evidence.

Theories make assertions that certain cause-effect relations are the case (cf. **[3] - [5]**), whereas hypotheses are causal accounts regarding past events leading to some set of observed effects (cf. **[4] - [5]**). Testing a theory or hypothesis requires that specific consequences be derived from them, such that if the theory or hypothesis is true, those consequences should be observed. The consequences sought in testing either theories or hypotheses are the same in that they would be effects stemming from particular causal conditions. But, beyond this, the nature of those effects are quite distinct, and as we will see later, it has been the failure to recognize this difference that has so often resulted in evolutionary biologists and ID proponents alike incorrectly characterizing the testing of their respective theories and hypotheses.

The formal structures for inferring potential test evidence of theories and hypotheses have the following respective forms:

[7] Potential test of a theory

Theory *T* to be tested:	given causal condition *x*, effect *y* will ensue
Experimental/causal conditions:	conditions *X*, *Y*, *Z*, ...*n* are to be established at time t_n, such that event *x* can occur
	---
Predicted consequence:	effect *y* should be observed at t_{n+1};

[8] Potential test of a hypothesis

Theory:	given causal condition *x*, effect *y* will ensue
Hypothesis *H* to be tested:	*x* occurred in the past
	--
Observations leading to original inference of *H* (cf. **[4]**, **[5]**):	effect *y*
Predicted consequences:	under conditions *X*, *Y*, *Z*, ...*n*, effects of type α, β, γ should be observed, that are related as specifically as possible to the causal conditions presented in *H*.

As indicated in **[7]**, the inference of potential tests of a theory stipulate conditions having the form of an experiment, where one must put themselves in a position of witnessing the 'experimental/causal conditions' as well as the subsequent 'predicted consequence.' The inference of possible tests of hypotheses are rather different in that the causal conditions occurred in the past. The hypothesis to be tested in **[8]** would have been the product of an earlier abductive inference (cf. **[3]**). An important consequence is that the effect(s) that prompted inference of the hypothesis in the first place could not then serve as test evidence of that hypothesis, much less could effects of the same class as those that led to the hypothesis. This latter implication stems from the fact that additional effects of the type that prompted the abductive inference of a hypothesis cannot be deduced from the premises in **[8]**. This limitation is immediately imposed by the nature of abduction. The inference of a hypothesis serves to provide at least some of the causal conditions that account for observed effects of a particular class. In other words, the hypothesis is only directed at the effects for which the hypothesis was inferred. It would not then be possible to deduce other effects of the type to which the hypothesis was originally directed. Since the hypothesis makes claims that specific causal events occurred in the past, then the only relevant potential test evidence that can be deduced would be effects that stem directly from those particular events. The implications of this requirement for both evolutionary biology and ID will be discussed below.

Inductive Inference: Testing Hypotheses and Theories

The examples in **[7]** and **[8]** only indicate the deductive derivation of potential observations that could serve as the evidence to test a theory or hypothesis, respectively. Simply inferring by way of deduction what one might observe is not tantamount to actually performing the test itself. This is a critical oversight in the characterization of science as being hypothetico-deductive. As we have seen thus far, hypothetico-deductivism ignores the mechanics that scientists apply to actually infer hypotheses and theories, while any focus on deduction only serves to point to what actions one might take in the subsequent event of actually engaging in testing. Testing is an inferential action that is inductive *sensu stricto* (cf. **[2]**).

From the deductions of potential tests in **[7]** and **[8]**, the formal structures of the tests themselves would have the following forms:

[9] Test of a theory

Theory *T* to be tested:	given causal condition *x*, effect *y* will ensue
Test conditions:	conditions *X*, *Y*, *Z*, ...*n* are established at time t_n, with causal event *x* is observed to take place
	--
Observed consequence:	effect of type *y* is observed at t_{n+1}; i.e., *T* is supported/ confirmed;

[10] Test of a hypothesis

Theory:	given causal condition *x*, effect *y* will ensue
Test conditions:	conditions *X*, *Y*, *Z*, ...*n* are established as the means to potentially observe consequences of type α, β, γ, given that *H* occurred
	--
Observed consequences:	effects of type α, β, γ are observed, i.e., supporting/ confirming that *H* represents actual past causal conditions.

Testing in both cases is inductive for the fact that when one does witness supporting or confirming test evidence, they have not established the truth of either a theory or hypothesis. The best that can be said is that positive test results allow continued acceptance until such time as disconfirming evidence is observed. Alternatively, in situations in which a theory or hypothesis fails critical tests, the inferences are also inductive in that the conclusions one draws are that there are other potential constructs that require future consideration.

What is especially important in the contexts of evolutionary biology and ID is whether or not tests of the form in **[9]** can be manifested for a particular theory. If so, then there is the opportunity to claim that empirical evidence is provided, either for or against that theory. As was pointed out in the previous section, and will be discussed in the next section in relation to ID, it is possible to deduce at least potential rudimentary test evidence for the presence of an intelligent agent as a causal factor. But actually carrying out the test as presented in **[9]** is not possible. This is the consequence that stands as the arbiter against ID.

Evolution and ID Cannot Be Defended Via Abductive Evidence

We can now to assess the popular claims that observed organisms, either extant or as fossil remains, provide evidence supporting any of the theories encompassed by evolutionary biology or ID. Representative arguments found in the evolutionary and ID literature will be examined next.

Failed Evidence For Evolution

The most consistent point of view regarding the testing of either a theory in evolutionary biology or a phylogenetic hypothesis has been that predictions of particular organisms, or their characters, qualify as relevant test evidence (e.g. Wiley 1975; Barnosky and Kraatz 2007; Prothero 2007; *contra* Fitzhugh 2006a, 2008a). In other words, the same class of evidence that is used to infer hypotheses, as well as a theory such as natural selection, also stands as supporting evidence. A good example of this position is the discovery of 'transitional forms' in the fossil record. Consider the recent description of *Tiktaalik roseae* by Daeschler et al. (2006; see also Shubin et al. 2006), a fossil of a Devonian fish with pectoral fins intermediate in form between fins and limbs. In commenting on this find, Ahlberg and Clack (2006: 748) regard the fossil as significant because "it demonstrates the predictive capacity of palaeontology." Similarly, *T. roseae* was cited in chapter one of the publication, *Science, Evolution, and Creationism* (National Academy of Sciences and Institute of Medicine 2008: 1, 3; see also Ayala 2006), within the section entitled, "The scientific evidence supporting biological evolution continues to grow at a rapid pace." Using this species as their example, it is claimed that,

> A prediction from more than a century of findings from evolutionary biology suggests that one of the early species that emerged from the Earth's oceans about 375 million years ago was the ancestor of amphibians, reptiles, dinosaurs, birds, and mammals. The discovery of *Tiktaalik* strongly supports that prediction. Indeed, the major bones in our own arms and legs are similar in overall configuration to those of *Tiktaalik.*
>
> The discovery of *Tiktaalik*, while critically important for confirming predictions of evolution theory, is just one example of the many findings made every year that add depth and breadth to the scientific understanding of biological evolution.

Is this fossil confirming evidence for "evolution theory?" Unfortunately, the answer is no. *Tiktaalik roseae* provides neither confirming evidence for an evolutionary transition from fishes to terrestrial tetrapods nor support for any of the theories within evolutionary biology. The following contrived example can make this evident.

A series of fossils are found from the dated strata shown in Figure 1A. The oldest fossils are of individuals lacking lateral appendages or dorsal ornamentation, whereas slightly younger fossils show the presence of one pair of appendages, and individuals from the youngest strata have three pairs. As well, members of the two youngest species have varying degrees of dorsal ornamentation. But, there is a gap in the record, between 165 and 100 mya where no members of this group had been found thus far. The phylogenetic hypothesis for these fossils (Fig. 1B) suggests the following two transformation series: (a) absence of appendages → one pair of appendages → three pairs of appendages, and (b) smooth dorsal margin → one pair of dorso-lateral protuberances → a row of five protuberances over the dorsum. The actual inference of these hypotheses from the observations would have the following abductive form (cf. Fitzhugh 2005a, b, 2006a, b, 2008a, b, in press):

[11] Theory: 'descent with modification' – if character *x*(0) exists among individuals of a reproductively isolated, gonochoristic or cross-fertilizing hermaphroditic population, and character *x*(1) originates by mechanisms *a*, *b*, *c*... *n*, and becomes fixed within the population by mechanisms *d*, *e*, *f*... *n* [= ancestral species hypothesis], followed by event or events *g*, *h*, *i*... *n*, wherein the population is divided into two or more reproductively isolated populations, then individuals to which descendant species hypotheses refer would exhibit *x*(1)

Observations (Fig. 1A):

(1) individuals to which species hypotheses *b-us*, *c-us*, and *d-us* refer have appendages in contrast to a convex ventrum as seen among individuals to which other species hypotheses refer;

(2) individuals to which species hypotheses *c-us* and *d-us* refer have dorsal protuberances in contrast to a convex dorsum as seen among individuals to which other species hypotheses refer

---

Hypothesis (Fig. 1B): (*a-us* (*b-us* (*c-us d-us*))), i.e.,

(1) ventral appendages originated by some unspecified mechanism(s) within a reproductively isolated population of individuals with a convex ventrum, and the condition became fixed in the population by some unspecified mechanism(s), followed by an unspecified event(s) that resulted in two reproductively isolated populations, i.e. members of species *a-us* and members of an ancestral species eventually leading to members of *b-us*, *c-us*, and *d-us*;

(2) dorsal protuberances originated by some unspecified mechanism(s) within a reproductively isolated population of individuals with a convex dorsum, and the condition became fixed in the population by some unspecified mechanism(s), followed by an unspecified event(s) that resulted in two reproductively isolated populations, i.e. members of species *c-us* and *d-us*.

From the temporal distribution (Fig. 1A) and phylogenetic hypothesis (Fig. 1B), one might claim that specimens will be found in the 165 to 100 mya strata that exhibit appendages and dorsal surface features that are transitional to what has been found in earlier and later strata (Fig. 1C). Such a prediction, if found to be the case, would be claimed to provide notable evidential support for the phylogenetic hypothesis, and by extension the ability of evolutionary biology to steer us in the right direction when it comes to seeking 'transitional forms.' Upon further investigations of the 165 to 100 mya strata fossils are found, with two pairs of appendages and three dorsal protuberances. Such a finding is in accord with what was 'predicted.' But, does this new fossil serve as a legitimate test of the phylogenetic hypothesis? Does the finding increase the veracity of evolutionary biology, at least in the context of, say, natural selection?

Let's first consider the situation if it is the case that fossils are test evidence for the phylogenetic hypothesis in **[11]** (Fig. 1B). To assert that the discovery of some 'transitional' form in strata between 165 and 100 mya supports (*a-us* (*b-us* (*c-us d-us*))), that prediction would have to be a deductive consequence of the form shown in **[8]**. The deduction would, at a minimum, have the following appearance:

[12a] Theory: 'descent with modification' – if character $x(0)$ exists among individuals of a reproductively isolated, gonochoristic or cross-fertilizing hermaphroditic population and character $x(1)$ originates by mechanisms a, b, c... n, and becomes fixed within the population by mechanisms d, e, f... n [= ancestral species hypothesis], followed by event or events g, h, i... n, wherein the population is divided into two or more reproductively isolated populations, then individuals to which descendant species hypotheses refer would exhibit $x(1)$

Hypothesis: (*a-us* (*b-us* (*c-us* *d-us*))), i.e.,

(1) ventral appendages originated by some unspecified mechanism(s) within a reproductively isolated population of individuals with a convex ventrum, and the condition became fixed in the population by some unspecified mechanism(s), followed by an unspecified event(s) that resulted in two reproductively isolated populations, i.e. members of species *a-us* and members of an ancestral species eventually leading to members of *b-us*, *c-us*, and *d-us*;

(2) dorsal protuberances originated by some unspecified mechanism(s) within a reproductively isolated population of individuals with a convex dorsum, and the condition became fixed in the population by some unspecified mechanism(s), followed by an unspecified event(s) that resulted in two reproductively isolated populations, i.e. members of species *c-us* and *d-us*

-

Observations that led to the original inference of the hypothesis (Fig. 1A):

(1) individuals to which species hypotheses *b-us* through *d-us* refer have appendages in contrast to a convex ventrum as seen among individuals to which other species hypotheses refer;

(2) individuals to which species hypotheses *c-us* and *d-us* refer have dorsal protuberances in contrast to a convex dorsum as seen among individuals to which other species hypotheses refer

'Predicted' test consequence, per Figure 1A:

> Fossils of a 'transitional' form should be found in strata between 165 and 100 mya that have a number of appendages or dorsal protuberances that are intermediate to what are seen among members of species *c-us* and *d-us* (cf. Fig. 1C).

What should be noticed in this example is that the deduction is invalid insofar as the conclusion of the 'predicted' test consequence is not possible. The premises, especially hypothesis (*a-us* (*b-us* (*c-us d-us*))), presents a series of causal conditions that, in conjunction with the theory of 'descent with modification,' account for the observed effects originally observed (Fig. 1A). From these premises there are no statements that could be derived that refer to specific properties of unobserved individuals. In point of fact, the premises do not preclude finding new fossils with *any possible combination* of characters. The consequence is that such fossils, if found, provide no relevant information that could assess the veracity of the hypothesis being tested. Additional fossil forms are simply newly observed effects also in need of explanation. The observations of new individuals cannot qualify as the test evidence for evaluating (*a-us* (*b-us* (*c-us d-us*))) (Fitzhugh 2005a, b, 2006a, b, 2008a, b). The only valid statements that could be inferred from the premises in **[12a]**, and would therefore function as potential tests of the hypothesis, would be those regarding effects that are direct and specific consequences of the causal events stated in the hypothesis, e.g.,

[12b] *Correct predicted consequences:*

> Respective effects *U*, *V*, *W*,... *n* should be observed, indicating the causal events of character origin and fixation of appendages and dorsal protuberances in an ancestral

population, and effects *X*, *Y*, *Z*,... *n* should be observed, indicating occurrences of causal events resulting in splittings of populations into separate, reproductively isolated groups.

Clearly, to speak of the testing of a phylogenetic hypothesis represented as a 'cladogram' (Fig. 1B, C) would not be possible given that such branching diagrams do not provide even sufficient specific details of the causal events one would be interested in testing. Inferring potential tests would only be possible once the different classes of events implied by such a diagram are fully explicated and replace the vaguely worded minor premise in **[12a]**.

One of the inherent difficulties with proceeding to the next step of actually testing explanatory hypotheses is that the process can be severely constrained by the amount of time that has elapsed between the hypothesized events to be tested and the observed effects. There is the possibility that the ensuing span of time is sufficient to eliminate the kind of test evidence identified in **[12b]**. In such a situation, we might say a hypothesis remains 'potentially testable,' as opposed to 'untestable.' In other words, empirical observations might eventually be available under current conditions. An untestable hypothesis would be one in which the consequences previously deduced are entirely beyond scrutiny because the theory or associated background knowledge upon which it is based, in the form of auxiliary theories, allow for no conceivable conditions under which the consequences can be empirically verified or refuted. Such an instance will be discussed in the next section (**Failed evidence for ID**).

A more general consideration regarding the observations in Figure 1A in terms of 'evidence' for any of the theories in evolutionary biology, such as natural selection, genetic drift, etc., suffers the problem just encountered for hypothesis testing. As in the instances described earlier of the fossils of *Tiktaalik roseae* being regarded as evidence supporting evolution, the failure of this position is that it rests on the evidence to which a theory would be applied, not evidence that could actually assess the theory. From the observations in Figure 1A, for one to claim that new fossils provide supporting evidence for an evolutionary theory, say natural selection, would mean that the deductive argument would have the form,

[13] Natural selection theory: in the event a novel, heritable character is derived as a consequence of mutation, and this character has a positive influence on fitness, then the character will become progressively prevalent in the population

'Causal' condition: fossils that are members of a new species should be found in strata between 165 and 100 mya

----

'*Predicted*' consequence: individuals will exhibit conditions intermediate of those observed in earlier and later individuals.

What is apparent in this instance is that the inference is not valid. The conjunction of the two premises would not allow for the conclusion. More fundamental, the minor premise does not state the proper antecedent conditions from which the theory of natural selection could be subjected to testing (cf. **[9]**). As theories have the form, 'given causal condition *x*, effect *y* will ensue,' deducing potential test observations would necessitate that the minor premise state the causal conditions that would need to be witnessed, from which consequences should be deducible if the theory is true. Testing theories proceed from knowing the causal conditions in the present to the subsequent observations of effects. Notice that this condition differs substantially from what was shown in **[10]** and **[12a/b]** for hypothesis testing, where causal conditions were past events that cannot be observed, such that what must be obtained is evidence related as specifically as possible to those events.

Rather than relying on fossils, which are the effects that have served as evidence for the abductive inferences of evolutionary theories and hypotheses, the proper inference of potential test evidence of the theory of natural selection would have the form,

[14]	**Natural selection theory:**	in the event a novel, heritable character is derived as a consequence of mutation, and this character has a positive influence on fitness, then the character will become progressively prevalent in the population
	Causal condition:	a new, heritable trait is observed within a population, that potentially enhances relative fitness
		---
	***Potential* Test consequence**:	occurrences of the trait should increase in the population into the future.

The evidential requirements for testing evolutionary hypotheses and theories immediately preclude simply recording the occurrences of fossils or extant organisms. What are observed of organisms are properties that prompt the applications of a variety of theories in evolutionary biology, such that those properties could not recursively assess the explanatory import of the theories and hypotheses abductively applied to them. Far more careful consideration of the nature of test evidence is required.

Failed Evidence For ID

The desire for ID advocates to establish their system of supernatural beliefs as a valid alternative to evolutionary biology, especially in terms of defending their position as one that should be considered in the realm of science, has mainly relied on two approaches: an ID theory is (1) testable, or (2) a concept derived from the principle of 'inference to the best explanation,' i.e. abduction. In both instances, the critical assessment of these approaches falls on whether or not the evidence being applied is appropriate to warrant judging any ID theory as amenable to scientific scrutiny in the manner to which theories in evolutionary biology have been subjected.

Is An ID Theory Testable? Claims by advocates that an ID theory is testable have not been particularly frequent, or consistent (Elsberry 2007). For instance, Dembski (1999: 258) states that, "Intelligent design is indeed testable, and it has been confirmed across a wide range of disciplines, spanning everything from natural history to molecular biology to information theory." Interestingly, however, Dembski provides no references to the actual research that substantiates this claim. Later, Dembski (2002a: 362) is somewhat more circumspect, especially when it comes to the requirement that testing must consider consequences deduced from observable antecedent conditions, as outlined earlier (cf. **[7]**, **[14]**), where,

> ...there is a sense in which to require prediction of intelligent design fundamentally misconstrues intelligent agency and design. To require of intelligent design that it predict specific instances of design in nature is to put design in the same boat as natural laws, locating their explanatory power in an

> extrapolation from past experience. This is to commit a category mistake. To be sure, designers, like natural laws, can behave predictably.... Yet unlike natural laws, which are universal and uniform, designers are also innovators. Innovation, the emergence to true novelty, eschews predictability. Designers are inventors. We cannot predict what an inventor would do short of becoming that inventor. Intelligent design presents a radically different problematic from a mechanistic science wedded solely to undirected natural causes. It offers predictability concerning the presence of design and the evolution of already existing designs, but it offers no predictability about fundamentally novel designs.

Dembski's (2002a) maneuver not only craftily attempts to distance ID from the rigors of testing, but it misconstrues the very mechanics of the testing process. The issue is not so much that of prediction, for that merely outlines specific consequences in the event one is willing to accept that some causal factor is the case (cf. **[7]**). Rather, the testing of any theory requires one to be in a situation to witness the conditions under which a causal agent is operative (cf. **[9]**). To engage in the testing of any theory, one must know as specifically as possible the causal parameters associated with the test, such that spurious conditions are minimized. Such requirements would already be outlined as part of the premises used to deduce predicted test consequences (cf. **[9]**). So it is not the matter of prediction that Dembski (2002a) wishes to avoid (e.g. **[7]**), but instead the reality that proper test conditions would be required. Notice as well that Dembski's (2002a) reference to natural laws is misplaced. Laws simply speak of the conjunctions of events, not the nature of causal relations. If, as Dembski would have us believe, that a supernatural designer is beyond reproach, then it would have to be acknowledged that there could be no test conditions under which one could evaluate the abilities of that designer to effect the origins of novel characters in organisms. Thus, Dembski (2002a: 366, emphasis original; see also Behe 1996: 196-197) does distance ID from the realm of critical scientific evaluation when he says,

> Proponents of intelligent design regard it as a scientific [*sic*] research program that investigates the effects of intelligent causes. Note that intelligent design studies the *effects* of intelligent causes and not intelligent causes *per se*. Intelligent design does not try to get into the head of a designing intelligence; rather, it looks at what a designing intelligence does and therewith draws inferences.

Reducing ID to the study of effects, to the exclusion of considering the actual causal conditions that produce those effects, removes the whole enterprise from actively seeking empirical evidence that is typically required to place a theory in good standing as a mechanism for acquiring understanding. Or to put the matter another way, the defense of ID is reduced to reliance on the effects being explained by the actions of a designer – it is the error of defending a theory on the basis of the 'evidence' used to abductively infer hypotheses based on an ID theory (Dawes 2007). This reduces the defense of ID to the circularity noted in the quote earlier from Popper (1992).

The inconsistencies and misconstrual of testing promulgated by Dembski (1999, 2002a) are even more evident in the treatment of 'irreducible complexity' by Michael J. Behe. In *Darwin's Black Box*, Behe (1996: 242, emphasis original) confuses the testing of explanatory hypotheses with the testing of theories:

> Hypothesis, careful testing, replicability – all these have served science well. But how can an intelligent designer be tested? Can a designer be put in a test tube? No, of course not. But neither can extinct common ancestors be put in test tubes. The problem is that whenever science tries to explain a unique historical event, careful testing and replicability are by definition impossible. Science may be able to study the motion of modern comets, and test Newton's laws of motion that describe how the comets move. But science will never be able to study the comet that putatively struck the earth many millions of years ago. Science can, however, observe the comet's lingering *effects* on the modern earth. Similarly, science can see the effects that a designer has had on life.

Beyond Behe's (1996) admission that an ID theory is not testable, let's be clear how he has misrepresented the very nature of testing. The assertion that a "unique historical event" cannot be tested is quite false. If Behe were correct, then all facets of science, indeed the process of progressive understanding, would be impossible. As has been outlined above, the mechanics of testing explanatory hypotheses is well established (cf. **[8]**, **[10]**). But, hypothesis testing is not the same as testing theories (cf. **[7]**, **[9]**). Without making this fundamental distinction, Behe (1996) has incorrectly characterized all aspects of testing in the historical and experimental sciences. The motivation for such misrepresentation is understandable, however, given that any ability to defend ID can only resort to the very effects ID is intended to explain. It is again an erroneous application of 'evidence' that might motivate one to consider ID as opposed to the proper 'evidence' necessary to support or refute ID. Behe (1996) provides no indication of the latter form.

In contrast to Behe's (1996) dismissal of testing, he presents an entirely different point of view in his response to reviews of his book (Behe 2001: 697, emphasis original; for a less coherent discussion of testing, see Behe 2002):

> In fact, *intelligent design is open to direct experimental rebuttal*. Here is a thought experiment that makes the point clear. In *Darwin's Black Box* I claimed that the bacterial flagellum was irreducibly complex and so required deliberate intelligent design. The flip side of this claim is that the flagellum can't be produced by natural selection acting on random mutation, or any other unintelligent process. To falsify such a claim, a scientist could go into the laboratory, place a bacterial species lacking a flagellum under some selective pressure (for mobility, say), grow it for ten thousand generations, and see if a flagellum – or *any* equally complex system – was produced. If that happened, my claims would be neatly disproven.

This is the only instance I am aware of where the experimental nature of testing has been characterized for ID. Unfortunately, Behe's (2001) test protocol is severely flawed. First, he does not correctly characterize the nature of "selective pressure." Selection, whether it is natural, sexual, or artificial, pertains to the properties of organisms. But such selection is only operative on novel characters at the point they are manifested as a consequence of mutation, intervention of a designer, etc. Selection allows for the opportunity for an advantageous character to become fixed in a population, not origin of the character. Alternatively, Behe's (2001) experimental setup could consider that a designer instantaneously converts all individuals in the culture to the flagellate condition.

Aside from the role of selection, Behe's (2001) experimental regime does not qualify as a valid test of the theory that a designer is engaged in the business of invoking the origin of novel characters. In point of fact, this concern is subsidiary to the more fundamental theory that a supernatural, intelligent agent even exists. This is a subject of concern that has been raised by Sober (2000, 2002, 2004, 2007) and Himma (2005). Namely, that any test of an ID theory requires consideration of the auxiliary theory of an intelligent agent. Consider the correct inferential structure for determining the test offered by Behe (2001) (cf. **[7]**):

[15]	Auxiliary theory:	a supernatural, intelligent agent exists
	Intelligent design theory:	a supernatural, intelligent agent has the capacity to invoke new features in living organisms
	Causal condition:	a culture of bacteria that lack flagella should be placed in the vicinity of a super-natural, intelligent agent at time t_n
		--
	Potential test consequence:	after 10,000 generations, i.e. at time t_{n+1}, at least some of the individuals in the culture should exhibit flagella.

Notice that articulating Behe's (2001) test in the proper form would require constraints confronted by any experiment: adequate exposition of the causal conditions. In this instance, for the test to be legitimate, the experimenter would at a minimum have to have empirical evidence of the occurrence of the intelligent agent in the vicinity of the culture. For this to be the case, however, it would already have been established by way of entirely separate test evidence that the auxiliary theory in **[15]** does hold. More on this in a moment. The test itself (cf. **[9]**) would appear as follows:

[16] Auxiliary theory:	a supernatural, intelligent agent exists
Intelligent design theory:	a supernatural, intelligent agent has the capacity to invoke new features in living organisms
Causal condition:	a culture of bacteria that lack flagella were established (Fig. 2A), and subsequently were placed in the vicinity of a supernatural, intelligent agent at time t_n (Fig. 2B)
	--
Confirming test consequence:	after 10,000 generations, i.e. at time t_{n+1}, at least some of the individuals in the culture exhibit flagella, *supporting the theory that a supernatural, intelligent agent has the ability to invoke novel characteristics in organisms* (Fig. 2D), or,
Disconfirming test consequence:	after 10,000 generations, i.e. at time t_{n+1}, no individuals in the culture exhibit flagella, *thus not supporting the theory that a supernatural, intelligent agent has the ability to invoke novel characteristics in organisms* (Fig. 2C).

Notice that performing the test in **[16]** requires that the experimenter witness, control, or otherwise have empirical evidence of the presence of the causal conditions. But, notice that manifesting those conditions is dependent upon the auxiliary theory that the intelligent agent does exist. This points to the specious nature of Behe's (2001) suggestion. Any attempt to actually test ID is not possible unless one can first establish through independent empirical evidence that there is a supernatural entity. An ID theory only presents the claim that an intelligent agent engages in a particular type of behavior, i.e. inducing novel characters in organisms. This theory is entailed by a more general theory of supernatural intelligent agents. In other words, to consider testing any ID theory already presumes that an auxiliary theory of supernatural intelligent agents has already been sufficiently tested and confirmed to warrant its extension into an ID theory. Ideally, the test in **[16]** is premature for the very fact that no tests have first been devised or conducted to determine the auxiliary theory (Sober 1999). If there is no empirical evidence for an intelligent agent, such that it could be considered as a causal factor in an experimental regime, then it makes no sense to speak of testing a theory wherein such an agent is a causal factor (cf. Pigliucci 2002: 129). This consideration was nicely summarized in the analysis by Himma (2005: 14, 25, respectively; see also Sober 2000, 2002, 2004, 2007), "What we lack in particular is some independent reason to think that there exists an intelligent agent with the appropriate abilities and motivations – which is, of course, exactly what the argument from specified biochemical complexity is trying to show," and "in their prototypical use, design inferences are used to identify intelligent behavior – and not to show the existence of the right kind of intelligent agents." In summary, Behe's (2001) experiment would have no ability to test ID since the most fundamental component of the theory, a designer, has not been established.

ID and 'Inference to the Best Explanation.' The subject of evidence capable of supporting ID has received greater attention among advocates in terms of the notion of explanatory power (e.g.

Dembski 1999; Meyer 1998b, 1999, 2002b, 2003, 2004b; Meyer et al. 2003). In this case, explanatory power is directly associated with abduction, under the phrase, 'inference to the best explanation.' The following statements encapsulate the ID position in this regard:

> ...the design inference from biological information constitutes an "inference to the best explanation." Recent work on the method of "inference to the best explanation"... suggests that determining which among a set of competing possible explanations constitutes the best depends upon assessments of the causal powers of competing explanatory entities. Causes that have the capability to produce the evidence in question constitute better explanations of that evidence than those that do not (Meyer 1998a: 138-139),
>
> Studies in the history and philosophy of science have shown that many scientific theories, particularly in the historical sciences, are formulated and justified as inferences to the best explanation.... Historical scientists, in particular, assess competing hypotheses by evaluating which hypothesis would, if true, provide the best explanation of some set of relevant data. Those with greater explanatory power are typically judged to be better – more probably true – theories (Meyer 2004a: 386),
>
> **inference to the best explanation:** a method of reasoning employed in the sciences in which scientists elect the hypothesis that would, if true, best explain the relevant evidence. Recent work in the philosophy of science has shown that

> those hypotheses that qualify as "best" typically provide simple, coherent, and causally adequate explanations of the evidence or phenomena in question. (Campbell and Meyer 2003: 618, glossary).

Meyer (2002b: 171; see also Meyer 1999, 2004b) goes so far as to claim that explanatory power is a surrogate for testing: "...theories may be evaluated and tested indirectly by the assessment of their explanatory power with respect to a variety of relevant data or 'classes of facts'." Dembski and Meyer (2002: 222-223, emphasis original; cf. Pennock 2007: 318) equate explanatory power with epistemic support:

> We believe an alternative understanding of epistemic support can foster a more productive interdisciplinary dialogue between science and theology. Fortunately such an alternative understanding is available. Although there are a number of ways to approach this alternative understanding of epistemic support, we approach it through the notion of *explanatory power*.

Dembski and Meyer (2002) then outline three criteria that enable one to determine a 'best explanation:' consonance, contribution, and comparative advantage. Consonance refers to the fit of a hypothesis to the effects inferred to account for those effects. Interestingly, Dembski and Meyer (2002b: 227) equate consonance with confirmation. Given hypothesis B and effects A, "if one were to take B as an (abductive) hypothesis, one would expect A to follow as a mater of course. To say B is consonant with A implies that A confirms B, where B is taken as a hypothesis." Contribution is claimed to be a corollary to the principle of simplicity, where a hypothesis "must perform some useful work in helping to explain" effects. Comparative advantage states that the better explanation is the one that out competes others.

What first needs to be recognized is that 'inference to the best explanation' can be interpreted in an abductive sense only insofar as one is speaking of the relations between theory and effects in the process of inferring a hypothesis (cf. **[3]** - **[6]**). Those relations provide no opportunity assess the veracity of the very theories that are applied in the inference. In the event a theory is applied as fully as possible to relevant effects, then the concluded hypothesis is 'best' from the standpoint that it makes the effects most probable given the truth of the hypothesis. But the sense in which this hypothesis is 'best' is only in the context of the theory applied to those effects. Consider again the abductive inferences in **[4]** and **[5]**, where the former applies the theory of natural selection and the latter the theory of ID to account for the same observations. Taken individually, each is an 'inference to the best explanation' in that they are successful at suggesting causal conditions. But what is at issue is the matter of which hypothesis is to be judged the more plausible, and the impetus for that concern comes not from the inferences themselves or the minor premises (effects in need of explanation) but from the alternate theories that are used. Questioning the alternate theories, as well as their respective auxiliary theories, stands entirely separate from the fact that they have been used in the two abductive inferences. Abductive inference or 'inference to the best explanation' is neither a tool for deciding among competing theories nor a surrogate for the evidence required to critically assess those theories. This implication has significant consequences for Dembski and Meyer's (2002a: 227) use of consonance, contribution, and comparative advantage. Consonance is immaterial unless one is engaging in inferences where the application of a theory to effects is not as complete as it should be, which simply translates into one not making apparent all relevant premises. This condition is often misrepresented in the phylogenetics literature as not correctly applying the principle of parsimony (Fitzhugh 2005a, 2006a, b). To say that consonance is equivalent to confirmation is erroneous. While the premises of any abductive inference are the very support for a conclusion, in terms of warranting the inference, those premises do not provide confirming evidence. Confirmation is an outcome of testing, i.e. induction *sensu stricto*, not abduction. Contribution has no relation to the principle of simplicity, which in the context of abduction is a comparative condition between one's causal questions and the ability of a hypothesis to answer those questions with as little extraneous information as possible (Sober 1975; Fitzhugh 2006a, b). Like consonance, comparative advantage is only relevant within particular abductive inferences, not comparisons of the theories used in those inferences.

Rather than consonance, contribution, and comparative advantage, the principle criteria used to select among competing theories, from which explanatory power can be referred, are consilience and simplicity (Thagard 1988; see Psillos 2002 for additional criteria). Consilience is a measure of the range of facts to which a theory can be applied. Obviously, one could claim that ID is more consilient than all of evolutionary biology. But consilience is mediated by simplicity. In the context of ID and a theory such as natural selection, the latter would be the simpler given that it does not require invoking at each instance a theory for which no independent empirical basis has yet to be provided (Pennock 2004). As noted earlier (see **Is an ID theory testable?**), the roadblocks to testing an ID theory are compounded by the fact that a more general theory of supernatural, intelligent agents remains untested. The consequence is that simplicity is immediately compromised each time ID is invoked relative to any theory in evolutionary biology.

Interestingly enough, the emphasis on explanatory power that has been popular among ID advocates is the very characteristic that shows the inadequacy of the theory. The explanatory power of a theory or hypothesis is determined by the extent to which it can increase the *a posteriori* probability of effects over the *a priori* probability (Rapoport 1972; Boyd 1985). As a consequence, assessing the explanatory power of theories and hypotheses comes from determining their predictive accuracy, i.e. the very activity eschewed by Dembski (2002a), quoted earlier (see **Is an ID theory testable?**). Thus, it seems legitimate that ID advocates would speak of initial explanatory power as a consequence of abduction, since the conjunction of some theory with effects will provide a hypothesis that makes those effects most probable. But the full measure of explanatory power must extend well beyond this inferential domain. Judging the explanatory power of a theory or hypothesis involves testing (e.g. **[9]** and **[10]**, respectively), which involves an entirely different form of evidence. It is the consequences of testing theories and hypotheses, i.e. their respective successes at making predictions, that is the mark of explanatory power. Abduction cannot stand as the appropriate mechanism, and thus cannot serve as a surrogate to testing (Lipton 1993, 2005). The 'evidence' that serves to abductively infer an ID hypothesis cannot stand as the 'evidence' upon which an ID theory can be judged.

Depew (2003: 446) offers a somewhat intermediate solution:

> Since [abduction] is the criterion that most intelligent design theorists wish to judge and be judged by, we may ask how the creationism-evolution debate stands when measured by this more reasonable, post-Popperian standard. Contemporary philosophers of biology who take a problem-oriented approach to science usually think of creationism as a genuine, but presently degenerative, research tradition. It is not to be dismissed because it is non-falsifiable, or even because it has been

> decisively falsified, but because its rate of problem-solving success has been declining steadily since its golden day.

As we have already seen, abduction is the beginning of the process of problem solving, not the end point. Thus, the only means to judge a theory 'degenerate' or lacking in problem-solving capability would come from the consequences that extend from the theory, and this is a matter of testing. A theory that is immune to testing is not so much degenerate as it is incapable of enhancing long-term understanding. It is the ability of an untestable theory to provide unassailable understanding that is degenerate.

Conclusion

The goal of science is to attain ongoing causal understanding – a state of affairs that is in constant flux. The present paper has attempted to show that allowing for growth of knowledge is contingent upon the elucidation of proper evidence from which evolutionary or ID theories should be judged. That evidence, supporting one theory over another, stands distinct from the evidence upon which tentative understanding is initially developed. Confusing the latter class of evidence for the former not only precludes constructive theory refinement, but also encourages the continued promotion of empirically baseless theories like ID.

To draw a distinction between science and non-science in the discussion of ID and evolutionary biology requires more than simply proclaiming any ID theory to be empirically baseless. Judging the merits of ID requires that scientists scrutinize such a theory, auxiliary theories, and associated hypotheses, relative to the different inferential procedures applied in any field of science, i.e. abduction, deduction, and induction. From such an analysis it is possible to identify what inferences are possible, such that the empirical success or failure of a given theory or hypothesis can be judged. Based on what has been discussed in this paper, a summary of what is and is not possible in the way of addressing evolutionary and ID theories and hypotheses is presented in Figure 3. In the case of both evolution and ID there are no restrictions on the abductive inferences to hypotheses and theories. The same holds for deductive inferences of potential test evidence. A rule that might, however, be imposed on both classes of inference is that they operate under the assumption that a theory and/or hypothesis is true, and that truth has already been affirmed as the result of testing. But such a rule might be too restrictive, especially given the importance of abduction to the production of theories, e.g. **[6]**. The class of inference that becomes especially critical is induction. In the case of testing a theory, what is required is an empirical record of the causal or experimental conditions from which test consequences are to ensue. And a significant limiting factor seen earlier is that auxiliary theories have themselves already been independently established through empirical evidence. For hypotheses, similar empirical requirements must be met in terms of consequences that have the lowest probability of occurrence if the hypothesized causal event(s) did not occur.

Figure 3 indicates the testing of a theory of ID is not possible, which is consistent with what was presented in **[16]**. Testing a theory of ID presumes that an intelligent agent can be witnessed under test conditions, whether such conditions are experimentally contrived or natural. With regard to the testing of ID-derived hypotheses, such testing would also be precluded for the fact that no empirical basis has been established for the auxiliary theory of a supernatural, intelligent agent that could function in a design capacity. As an ID-based hypothesis has as it's foundation an ID theory that is not amenable to actual testing, then by extension the hypothesis is also immune to testing. But, as indicated in Figure 3, there are some ID/creationism hypotheses that can be, and have been tested (Prothero 2007), but there are important caveats to be recognized. Consider for instance hypotheses that the universe and Earth are no older than 6,000 years, or that the Noachian flood occurred. What is significant to notice is that each has been tested in the context of observing evidence refuting the factual claims they present. For instance, hypotheses of the form, 'God created the universe and Earth 6,000 years ago' and 'God caused the Earth to be covered by a massive flood' make very specific causal assertions. While components of these hypotheses have been refuted, i.e. ages of the universe and Earth, and the occurrence of a global flood, those refutations cannot extend to the claims of supernatural powers included in the hypotheses. Thus, the 'not possible' and 'possible' conditions for ID hypotheses in Figure 3 pertain to the duality of non-empirical and empirical claims that cannot and can be tested, respectively. As with any theory founded on supernatural causes, a supernatural-based hypothesis will not be available for testing from the standpoint of actually evaluating the causal conditions it offers.

The different contexts in which evidence plays a role in the process of acquiring understanding suggests that a demarcation criterion is possible based on whether or not a theory or hypothesis can be subjected to the empirical scrutiny of testing. But, this limitation on testing does not necessarily preclude one from engaging in the abductive inferences of ID-based theories and hypotheses, or even entertaining deductions to potential test consequences. It is only when consideration of how evidence is applied in the different classes of inference that science can be used to evaluate the relative standing of theories in evolutionary biology and ID, as useful tool for the acquisition of understanding. As noted by Sober (1999: 68; see also

Fitelson et al. 1999), "There is more to science than the activity of running tests. Yet, I think that the scientific enterprise is directed towards the goal of bringing problems to an empirical resolution, or setting them aside when this cannot be done. Science, I am suggesting, is the art of the testable."

References

Achinstein P. Inference to scientific laws. In: Stuewer RH editor. Volume V: historical and philosophical perspectives of science, Minnesota studies in the philosophy of science. Minneapolis: University of Minnesota Press; 1970. p. 87-111.

Ahlberg E, Clack JA. A firm step from water to land. Nature. 2006;440:747-749.

Atocha A. Abductive reasoning: logical investigations into discovery and explanation. Dordrecht, The Netherlands: Springer; 2006.

Ayala F. Design without designer: Darwin's greatest discovery. In: Dembski WA, Ruse M editors. Debating design: from Darwin to DNA. New York: Cambridge University Press; 2004. p. 55-80.

Ayala FJ. Darwin and intelligent design. Minneapolis, MN: Fortress Press; 2006.

Barnosky AD, Kraatz BR. The role of climatic change in the evolution of mammals. BioScience. 2007;57:523-532.

Behe MJ. Reply to my critics: a response to reviews of Darwin's Black Box: The Biochemical Challenge to Evolution. Biol Phil. 2001;16:685-709.

Behe MJ. Answering scientific criticisms of intelligent design. In: Behe MJ, Dembski WA, Meyer SC editors. Science and evidence for design in the universe. San Francisco, CA:

Ignatius Press; 2002. p. 133-149.

Behe MJ. Darwin's black box: the biochemical challenge to evolution. New York: Free Press; 2006.

Ben-Menahem Y. The inference to the best explanation. Erkenntnis. 1990;33:319-344.

Boyd R. Observations, explanatory power, and simplicity: toward a non-Humean account. In: Achinstein P, Hannaway O editors. Observation, experiment, and hypothesis in modern physical science. Cambridge, MA: MIT Press; 1985. p. 47-94.

Campbell JA, Meyer SC editors. Darwinism, design, and public education. East Lansing: Michigan State University Press; 2003.

Charlesworth B, Charlesworth D. Evolution: a very short introduction. New York: Oxford University Press; 2003.

Copi IM, Cohen C. Logic. Upper Saddle River, NJ: Prentice Hall; 1998.

Curd MV. The logic of discovery: an analysis of three approaches. In: Nickles T editor. Scientific discovery, logic and rationality. Dordrecht, The Netherlands: D. Reidel Publishing Company; 1980. p. 201-219.

Daeschler EB, Shubin NH, Jenkins Jr. FA. A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. Nature. 2006;440:757-763.

Darwin CR. The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. London: John Murray; 1872.

Davis P, Kenyon DH. Of pandas and people: the central question of biological origins. Dallas, TX: Haughton Publishing Company; 2004.

Dawes GW. What *is* wrong with intelligent design? Int J Philos Relig. 2007;61:69-81.

Dembski WA. The design inference – eliminating chance through small probabilities. Cambridge, MA: Cambridge University Press; 1998.

Dembski WA. Intelligent design: the bridge between science & theology. Downers Grove, IL: InterVarsity Press; 1999.

Dembski WA. No free lunch: why specified complexity cannot be purchased without intelligence. New York: Rowman & Littlefield Publishers, Inc.; 2002a.

Dembski WA. The third mode of explanation: detecting evidence of intelligent design in the sciences. In: Behe MJ, Dembski WA, Meyer SC editors. Science and evidence for design in the universe. San Francisco, CA: Ignatius Press; 2002b. p. 17-51.

Dembski WA, Meyer SC. Fruitful interchange or polite chitchat? The dialogue between science and theology. In: Behe MJ, Dembski WA, Meyer SC editors. Science and evidence for design in the universe. San Francisco, CA: Ignatius Press; 2002. p. 213-234.

Dembski WA, Wells J. The design of life: discovering signs of intelligence in biological systems. Dallas, TX: The Foundation for Thought and Ethics; 2008.

Depew D. Intelligent design and irreducible complexity: a rejoinder. In: Campbell JA, Meyer SC editors. Darwinism, design, and public education. East Lansing: Michigan State University Press; 2003. p. 441-454.

Douven I. Testing inference to the best explanation. Synthese. 2002;130:355-377.

Endler JA. Natural selection in the wild. Princeton, NJ: Princeton University Press; 1986.

Elsberry WR. 2007. Logic and math turn to smoke and mirrors: William Dembski's "design inference." In: Petto AJ, Godfrey LR editors. Scientists confront intelligent design and creationism. New York: W.W. Norton & Company; 2007. p. 250-271.

Fann KT. Peirce's theory of abduction. The Hague, The Netherlands: Martinus Nijhoff; 1970.

Fitelson B, Stephens C, Sober E. How not to detect design – critical notice: William A. Dembski, The Design Inference. Phil Sci. 1999;66:472-488.

Fitzhugh K. Les bases philosophiques de l'inférence phylogénétique: une vue d'ensemble. Biosystema. 2005a;24:83-105.

Fitzhugh K. The inferential basis of species hypotheses: the solution to defining the term 'species.' Mar Ecol. 2005b;26:155-165.

Fitzhugh K. The abduction of phylogenetic hypotheses. Zootaxa. 2006a;1145:1-110.

Fitzhugh K. The 'requirement of total evidence' and its role in phylogenetic systematics. Biol Phil. 2006b;21:309-351.

Fitzhugh K. The philosophical basis of character coding for the inference of phylogenetic hypotheses. Zool Scr. 2006c;35:261-286.

Fitzhugh K. Fact, theory, test and evolution. Zool Scr. 2008a;37:109-113.

Fitzhugh K. Abductive inference: implications for 'Linnean' and 'phylogenetic' approaches for representing biological systematization. Evol Biol. 2008b; 35:52-82.

Fitzhugh K. Clarifying the role of character loss in phylogenetic inference. Zool Scr. in press.

Futuyma DJ. Evolutionary biology. Sunderland, MA: Sinauer Associates, Inc.; 1998.

Futuyma DJ. Evolution. Sunderland, MA: Sinauer Associates, Inc.; 2005.

Gauch Jr. HG. Scientific method in practice. New York: Cambridge University Press; 2003.

Godfrey-Smith P. Theory and reality: an introduction to the philosophy of science. Chicago, IL: University of Chicago Press; 2003.

Hacking I. An introduction to probability and inductive logic. New York: Cambridge University Press; 2001.

Hanson NR. Patterns of discovery: an inquiry into the conceptual foundations of science. New York: Cambridge University Press; 1958.

Harman G. The inference to the best explanation. Phil Rev. 1965;74:88-95.

Hausman DM. Causal asymmetries. New York: Cambridge University Press; 1998.

Hempel CG. Aspects of scientific explanation and other essays in the philosophy of science. New York: The Free Press; 1965.

Himma KE. The application-conditions for design inferences: why the design arguments need the help of other arguments for God's existence. Int J Philos Relig. 2005;57:1-33.

Josephson JR, Josephson SG editors. Abductive inference: computation, philosophy, technology. New York: Cambridge University Press; 1994.

Kingsolver JG, Pfennig DW. Patterns and power of phenotypic selection in nature. BioScience. 2007;57:561-572.

Lankester ER. On the use of the term homology in modern zoology, and the distinction between homogenetic and homoplastic agreements. Ann Mag Nat Hist. 1870;6:35-43.

Laudan L. Commentary: science at the bar – causes for concern. Sci Tech Human Val. 1982;7:16-19.

Laudan L. The demise of the demarcation problem. In: Cohen RS, Laudan L. editors. Physics, philosophy, and psychoanalysis. D. Dordrecht, The Netherlands: D. Reidel Publishing Company; 1983. p. 111-127.

Magnani L. Abduction, reason, and science: processes of discovery and explanation. New

York: Kluwer Academic; 2001.

Lipton P. Inference to the best explanation. New York: Routledge; 1993.

Lipton P. Testing hypotheses: prediction and prejudice. Science. 2005;307:219-221.

Mahner M, Bunge M. Foundations of biophilosophy. New York: Springer; 1997.

Meyer SC. Laws, causes, and facts: response to Michael Ruse. In: Buell J, Hearn V editors. Darwinism: science or philosophy? Richardson, TX: Foundation for Thought and Ethics; 1994. p. 29-40.

Meyer SC. The explanatory power of design: DNA and the origin of information. In: Dembski WA editor. Mere creation: science, faith & intelligent design. Downers Grove, IL: InterVarsity Press; 1998a. p. 113-147.

Meyer SC. DNA by design: an inference to the best explanation for the origin of biological information. Rhet Pub Affairs. 1998b;1:519-556.

Meyer SC. The return of the God hypothesis. J Interdisc Stud. 1999;11:1-38.

Meyer SC. Evidence for design in physics and biology: from the origin of the universe to the origin of life. In: Behe MJ, Dembski WA, Meyer SC editors. Science and evidence for design in the universe. San Francisco, CA: Ignatius Press; 2002a. p. 53-111.

Meyer SC. The scientific status of intelligent design: the methodological equivalence of naturalistic and non-naturalistic origins theories. In: Behe MJ, Dembski WA, Meyer SC editors. Science and evidence for design in the universe. San Francisco: Ignatius Press; 2002b. p. 151-211.

Meyer SC. DNA and the origin of life: information, specification, and explanation. In: Campbell JA, Meyer SC editors. *Darwinism, design, and public education*. East

Lansing: Michigan State University Press; 2003. p. 223-285.

Meyer SC. The Cambrian information explosion: evidence for intelligent design. In: Dembski WA, Ruse M editors. *Debating design: from Darwin to DNA*. New York: Cambridge University Press; 2004a. p. 371-391.

Meyer SC. The origin of biological information and the higher taxonomic categories. Proc Biol Soc Wash. 2004b;117:213-239.

Meyer SC, Keas MN. The meanings of evolution. In: Campbell JA, Meyer SC editors. *Darwinism, design, and public education*. East Lansing: Michigan State University Press; 2003. p. 135-156.

Meyer SC, Ross M, Nelson P, Chien P. The Cambrian explosion: biology's big bang. In: Campbell JA, Meyer SC editors. *Darwinism, design, and public Education*. East Lansing: Michigan State University Press; 2003. p. 323-402.

Millman AB, Smith CL. Darwin's use of analogical reasoning in theory construction. Metaphor Symbol. 1997;12:159-187.

Mitton JB. Selection in natural populations. New York: Oxford University Press; 1997.

National Academy of Sciences and Institute of Medicine. Science, evolution, and creationism. Washington DC: The National Academies Press; 2008.

Nickles T. Introductory essay: scientific discovery and the future of philosophy of science. In: Nickles T editor. Scientific discovery, logic and rationality. Dordrecht, The Netherlands: D. Reidel Publishing Company; 1980. p. 1-59.

Peirce CS. Illustrations of the logic of science. Sixth paper. – Deduction, induction, and hypothesis. Pop Sci Month. 1878;13:470-482.

Peirce CS. Collected papers of Charles Sanders Peirce, volumes 1-6. Hartshorne C, Weiss P, Burks A editors. Cambridge, MA: Harvard University Press; 1931-1935.

Peirce CS. Collected papers of Charles Sanders Peirce, volumes 7-8. Burks A editor. Cambridge, MA: Harvard University Press; 1958.

Pennock RT. DNA by design? Stephen Meyer and the return of the God hypothesis. In: Dembski WA, Ruse M editors. Debating design: from Darwin to DNA. New York: Cambridge University Press; 2004. p. 130-148.

Pennock RT. God of the gaps: the argument from ignorance and the limits of methodological naturalism. In: Petto AJ, Godfrey LR editors. Scientists confront intelligent design and creationism. New York: WW Norton & Company; 2007. p. 309-338.

Pigliucci M. Denying evolution: creationism, scientism, and the nature of science. Sunderland, MA: Sinauer Associates, Publishers; 2002.

Popper KR. Objective knowledge: an evolutionary approach. New York: Oxford University Press; 1983.

Popper KR. Realism and the aim of science. New York: Routledge; 1992.

Prothero DR. Evolution: what the fossils say and why it matters. New York: Columbia University Press; 2007.

Psillos S. Simply the best: a case for abduction. In: Kakas AC, Sadri F editors. Computational logic: logic programming and beyond. New York: Springer; 2002. p. 605-625.

Rapoport A. Explanatory power and explanatory appeal of theories. Synthese. 1972;24:321-342.

Reilly FE. Charles Peirce's theory of scientific method. New York: Fordham University Press; 1970.

Rescher N. Scientific explanation. New York: The Free Press; 1970.

Salmon WC. Scientific explanation and the causal structure of the world. Princeton, NJ: Princeton University Press; 1984a.

Salmon WC. Logic. Englewood Cliffs, NJ: Prentice-Hall, Inc.; 1984b.

Scott EC. Evolution vs. creationism: an introduction. Los Angeles, CA: University of California Press; 2004.

Shubin NH, Daeschler EB, Jenkins FA. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. Nature. 2006;440:764-771.

Sober E. Simplicity. New York: Oxford University Press; 1975.

Sober E. Testability. Proc Addresses APA. 1999;73:47-76.

Sober E. Philosophy of biology. Boulder, CO: Westview Press; 2000.

Sober E. Intelligent design and probability reasoning. Int J Philos Relig. 2002;52:65-80.

Sober E. The design argument. In: Dembski WA, Ruse M editors. Debating design: from Darwin to DNA. New York: Cambridge University Press; 2004. p. 98-129.

Sober E. What is wrong with intelligent design? Q Rev Biol. 2007;82:3-8.

Stamos DN. Popper, laws, and the exclusion of biology from genuine science. Acta Biotheor. 2007;55:357-375.

Thagard P. Computational philosophy of science. Cambridge, MA: The MIT Press; 1988.

Van Fraassen BC. The scientific image. New York: Clarendon Press; 1990.

Walton D. Abductive reasoning. Tuscaloosa: The University of Alabama Press; 2004.

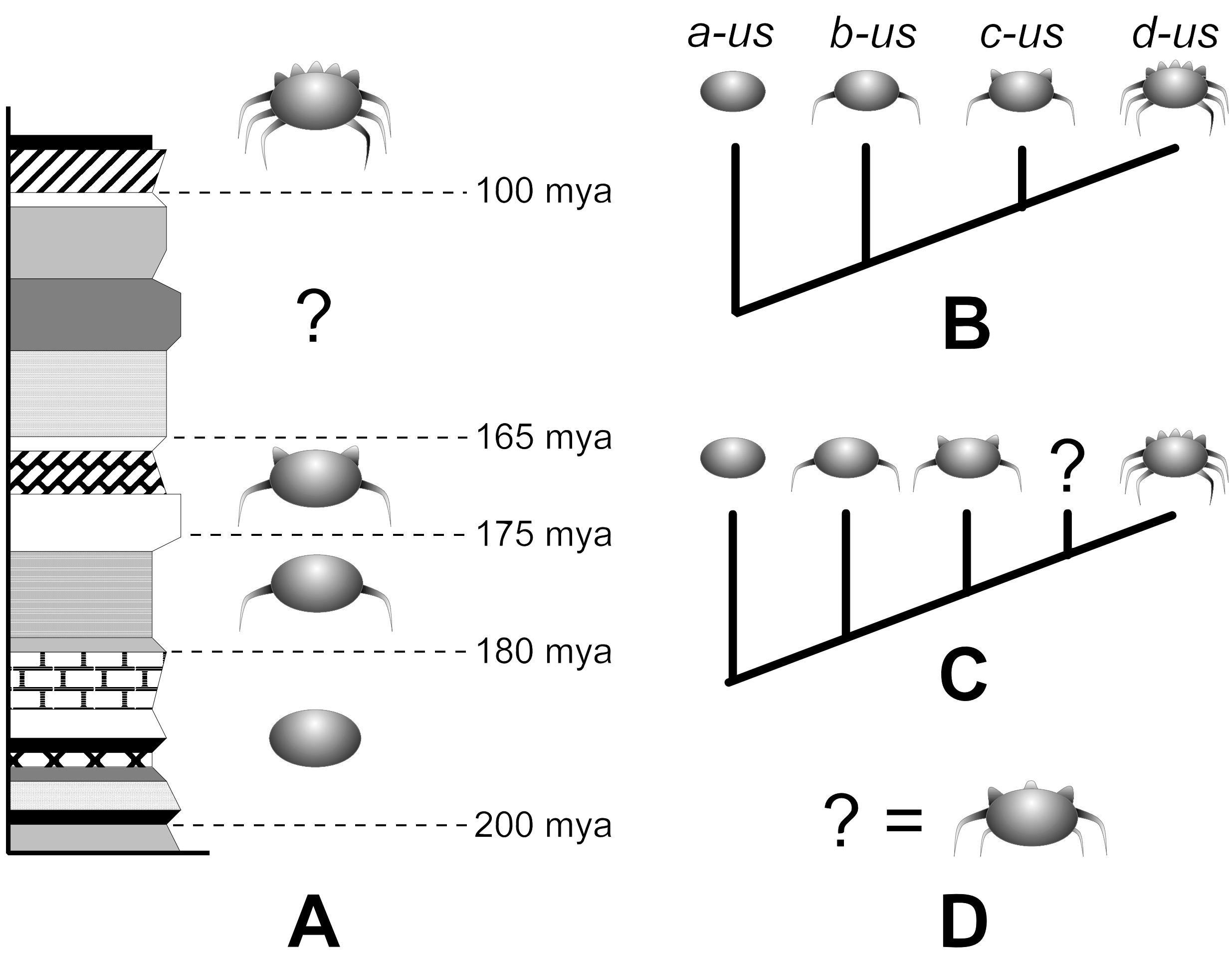
Wiley EO. Karl R. Popper, systematics, and classification: a reply to Walter Bock and other evolutionary taxonomists. Syst Zool. 1975;24:233-243.

Figure legends

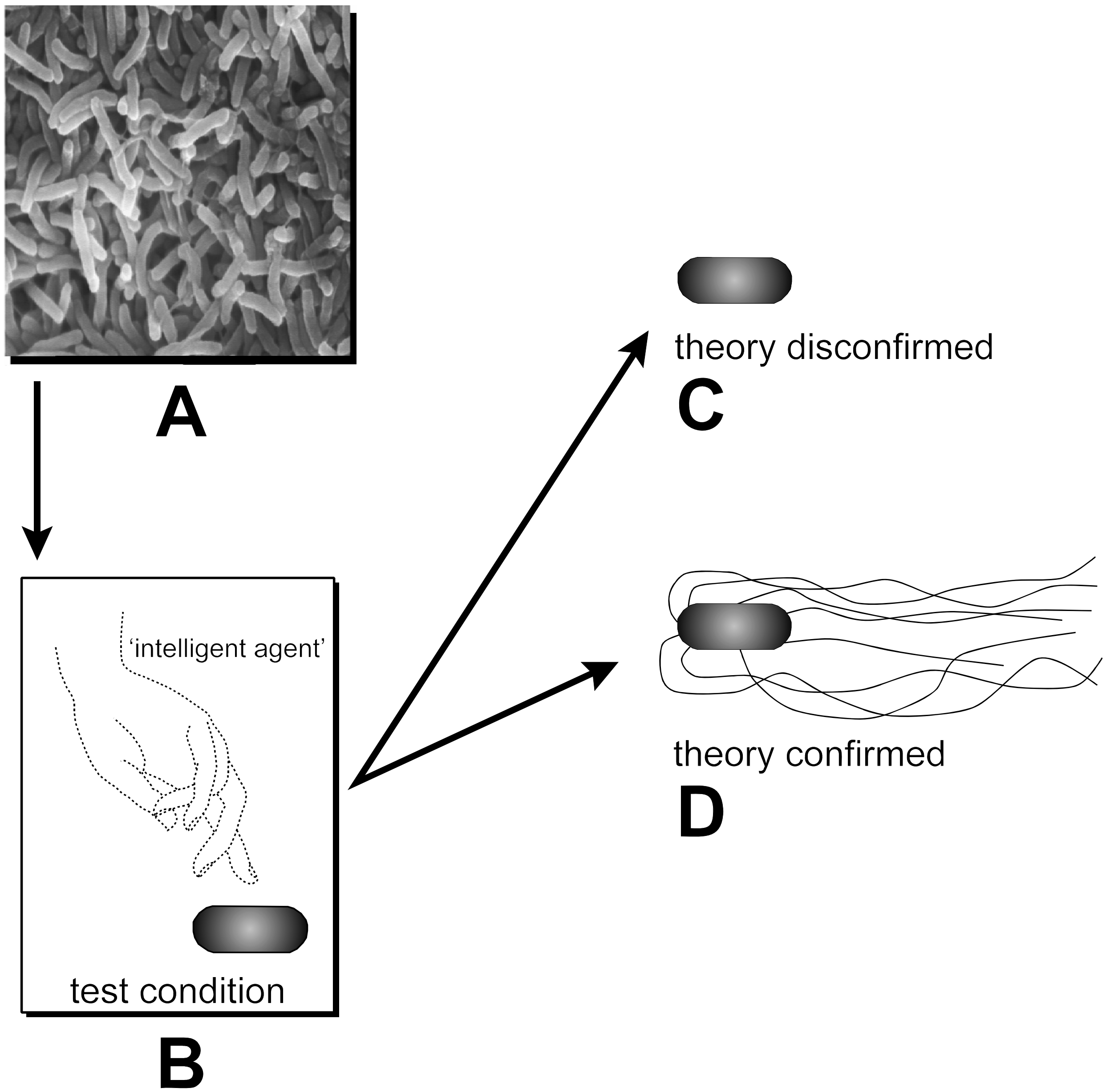
Fig 1 example of the use of 'transitional' fossils as evidence of evolution. **A.** Occurrences of fossils, implying evolutionary transformations in appendages and dorsal surface ornamentation. **B.** A phylogenetic hypothesis, providing explanatory accounts for characters observed among the fossils. **C.** Given the absence of fossils in the 165 to 100 mya strata, it might be 'predicted' that a transitional form, indicated by '?', might eventually be found that confirms not only some principle of evolution, as well as the phylogenetic hypothesis. **D.** The 'transitional' fossil observed subsequently found in 165 to 100 mya strata. See text for a discussion of why such a fossil does not provide confirming evidence for either the principle of evolution or the hypothesis in **A**

Fig 2 The formal experimental protocol needed to test the theory of ID, based on the suggestion by Behe (2001). See text for discussion

Fig 3 Summary of the inferential procedures one might encounter when applying either the theory of ID or theories in evolutionary biology. Given that abduction and deduction allow for the respective activities of introducing theories and hypotheses, and suggestions of what test consequences one should seek, it does not appear that science places any constraints on those activities. While the inferences of theories/hypotheses and deductive consequences might not be constrained, the critical demarcation is that science requires that actual (or even potential) testing be possible. A theory of ID, and certain hypotheses, are not amenable to testing for the fact that the requisite causal conditions cannot be manifested or the auxiliary theory of an intelligent agent has not been established by independent evidence. See text for discussion



Fitzhugh -- Fig. 1



Fitzhugh -- Fig. 2

	Intelligent Design		Evolutionary Biology	
	Theories	Hypotheses	Theories	Hypotheses
Abduction	possible	possible	possible	possible
Deduction	possible	possible	possible	possible
Induction	not possible	not possible / possible[1]	possible	possible

[1]Testing ID/creationism hypotheses are extremely limited in scope. Such instances would only address the claim that some physical event occurred, not the actual supernatural cause of the event, e.g. the Earth is less than 6,000 years old or that the Noachian deluge occurred.

Fitzhugh -- Fig. 3