A number of factors contribute to the success of such an approach. The first is that the predictions are multi-cornered such that it is clearly specified what sort of evidence would adjudicate between the hypotheses. Secondly, the terms 'allopatric' and 'sympatric' are clearly distinguished in genetic terms although this is implicit. Thirdly, there is implicit recognition of the historical aspect of speciation. By the principle of uniformitarianism the processes or stages that exist today are taken to be representative of the past. Fourthly, the right question is being asked. The important question is not which mode of speciation is possible, but which is the prevailing and hence most important mode. Fifthly, Williams' predictions are at the right level. They are not about individual cases from which we may only learn whether or not a particular mode has occurred. Instead they are about patterns in the distributions of a group of species pairs.

What we can do now is to construct the entire argument. We will only consider allopatric and sympatric speciation here, but a similar table could be constructed to compare allopatric speciation with other modes.

Table 5-4. Multi-cornered predictions between allopatric and symaptric speciation.

		the conditions required for allopatric speciation are likely to occur.
If allopatric speciation is the prevailing mode	then	there will be a correlation between speciation and the presence of geographic barriers.
		speciation will occur most frequently in groups with poor methods of dispersal.
		etc.
		the conditions required for sympatric speciation are likely to occur.
If sympatric speciation is the prevailing mode	then	there will be little or no correlation between speciation and the presence of geographic barriers.
		speciation will occur most frequently in groups with good methods of dispersal.
		etc.

As we shall see, contrasting the predictions of the various modes is essential if we are to correctly practice the next step- the assessing of the evidence.

Step 4: Adjudicating between alternatives

The object of the previous step was to make multi-cornered deductions from various modes of speciation. In this step the object is to come to a decision about the adequacy of these modes from the evidence. Because our focus is on methodology we will not consider the evidence for and against all modes of speciation. Discussion will be restricted to allopatric and sympatric speciation.

Now, the fact that the field of speciation is known for its pluralism might suggest , *prima facie*, that the evidence needed to make decisions is not forthcoming. I hope to show that the problem is not primarily a lack of evidence.

To the question 'does allopatric speciation ever occur' the answer is unequivocal. It is yes. There are many cases where the evidence points, and only points to allopatric speciation. This is generally accepted so we will not go into any detail here. Mayr (1963:481-515) has reviewed some of the evidence. Some of the more classic work has been done on the Drosophilidae on the Hawaiian islands (Carson 1975), finches on the Galapagos islands (Maynard-Smith 1975), *Zosterops* (a land bird) in the Central Solomon islands (Mayr 1942:227), and ant thrushes in South America (Mayr 1942:228) to mention just a few cases.

In contrast the evidence for sympatric speciation is very weak. As we have seen the classic proofs of sympatric speciation turn out, on closer examination, to be accountable in terms of other modes (Mayr 1963, Futuyma and Mayer 1980, Templeton 1981, Paterson 1981).

To the question 'which is the most likely mode of speciation' the answer is again unequivocal. It is allopatric speciation. The sorts of conditions required for allopatric speciation occur commonly in nature. Templeton (1981), Paterson (1978, 1981), Masters and Paterson (1987) have reviewed some of the evidence. These authors have noted, on the other hand, that the conditions required for sympatric speciation are extremely restrictive so that this mode is unlikely to ever occur.

To the question which is the most prevalent mode of speciation the answer is again unequivocal. It is allopatric speciation. Distributional studies have confirmed this. Most incipient species pairs live outside of each other's dispersal range (Mayr 1963, Ghiselin 1969, Williams 1973).

Even with all this data, sympatric speciation has enjoyed a favour that is out of proportion to the evidence for it. White for instance says that there can be no doubt about the existence of sympatric speciation, at least in many group of insects. The problem, then, is not primarily an empirical one. This is not to say that that we could not do with more evidence. Masters and Paterson (1987) have pointed out areas in which more research is needed. However, it would be a mistake to think that pluralism could be avoided by simply appealing to the evidence.

No, the problem is at least to some extent a philosophical one. It concerns our approach to evidence. If acceptance of a mode is based simply on the existence of some positive evidence for it, and the limited value of common evidence is not recognised for what it is (and we have seen this happen), then a chain of unfortunate consequences will follow. Our standards will be low. It will be enough that a mode is conceivable given the available evidence. Since common evidence is easy to come by, many modes will be conceivable. As a result there will seem to be no need to develop multi-cornered predictions. This in turn will mean that the discriminatory evidence that does exist will not be recognised for what it is. Next, the different empirical credentials of the various modes will not be appreciated and this is exactly what Paterson (1981) has reported.

He has seen the need to distinguish between Class I and Class II models. Class I models are theoretical with "support from observed critical facts", whereas Class II models are "without support from observed critical facts". The criterion for the acceptance of a model is the existence of "reliable, critical, observational data" which "compel acceptance of the view that speciation in accordance with this mode has actually occurred."

The final consequence is pluralism and a pluralism that exists not because speciation occurs in many ways, but because our standards of acceptance are low. Paterson reported a need "for a more consistent logic if we are to reduce disputation and disagreement in Evolutionary Biology." It is submitted that we will go a long way to achieving this by carefully distinguishing between the different kinds of evidence.

CHAPTER 6: DEMARCATION AND METHOD IN A SIMULATION GAME

Mathematical games have been used to simulate the scientific enterprise and to teach scientific method (Gardner 1977,1979). One of these, which will be called Target, is quite a good analogy of science. It also goes by the name Bulls and Cows. A variant of the game has been commercially packaged as Mastermind and computer programmes have been written for both games (Ahl 1979:96-98). If Target is played properly it mimics the development of science. If not it mimics bad science or pseudo-science. As a result it has value as a heuristic device; suggesting solutions in both the areas of demarcation and method. As with all analogies, Target has similarities and dissimilarities with what it represents. We shall see that it is similar to science at places where the otherwise fruitful evolutionary metaphor of the growth of scientific knowledge (Toulmin 1967, 1970:46-47, Kuhn 1962:146, 1970:264, Popper 1972a:108, 1979:261-265) breaks down. According to evolutionary epistemology there is a "natural selection" of hypotheses. Rival hypotheses of lower fitness are eliminated. If one holds to the view that science is purposeful and progressive, then this analogy breaks down because evolution is blind (Ruse 1983:136).We shall see that in this respect the target analogy correlates more closely with science.

A second advantage is that the "theories" and "observations" of the game involve numbers which are neutral with respect to our prior commitments. Methodological points can be illustrated without the complication of possible prejudices. The chapter will take the following course:

Firstly, the rules of the game will be described. Then I will show how the game illustrates principles of demarcation and method. The similarities and dissimilarities between the game and science will be discussed. I will show that the game illuminates certain distinctions made by many philosophers of science. Finally I will show how the game can be used as a didactic tool.

How to play Target

Target is played by two people. The object is for player 'A' to arrive at a four digit number which is kept hidden from him by player 'B'. This number is called the target number. 'A' makes repeated calls of four digit numbers, and using his logic, and the strictly defined responses of 'B', tries to discover the target number. 'B' responds to 'A's calls by giving the number of 'bulls' and 'hits'. One bull means that one digit in 'A's call is in the target number and in the right place. One hit means that one digit is in the target number but in the wrong place. Figure 6-1 gives a few steps in the game as an example.

Target Number

7605

A's Calls		B's Responses
No. 1	0123	1 Hit
No. 2	6123	1 Hit
No. 3	6023	2 Hits
No. 4	6302	1 Bull, 1 Hit

Figure 6-1. A few steps in target.

Similarities between Target and science

The features of the game which correspond to the scientific enterprise are the following:

1. **The Target number**. This number represents the real world or a particular part of it. If player "B" chooses the call then the number exists independently of player "A."

2. **The calls**. These represent hypotheses or theories. They are guesses of the target number.

3. **The responses**. These represent observations.

Target and the problem of demarcation

Previously a fundamental statement was given concerning the nature of science, that theories in scientific disciplines are constrained by the real world. This feature is paralleled in the game if it is played properly. The calls made are constrained by the target number. Given this feature we can predict that the game should have certain properties. It would have the properties of a negative feedback system (see fig. 6-2). Science has often been described as such a system (e.g. Medawar and Medawar 1977:15- 16, Ṡparkes 1981). The target number limits the range of acceptable calls. At any given step after the first call there is more than one appropriate call, but there are clearly unacceptable calls. What determines the acceptability of calls is the target number, through the responses to each call. This negative feedback property leads to a progressive limiting of the range of acceptable calls. Thus the game, like science, shows exclusionary convergence.

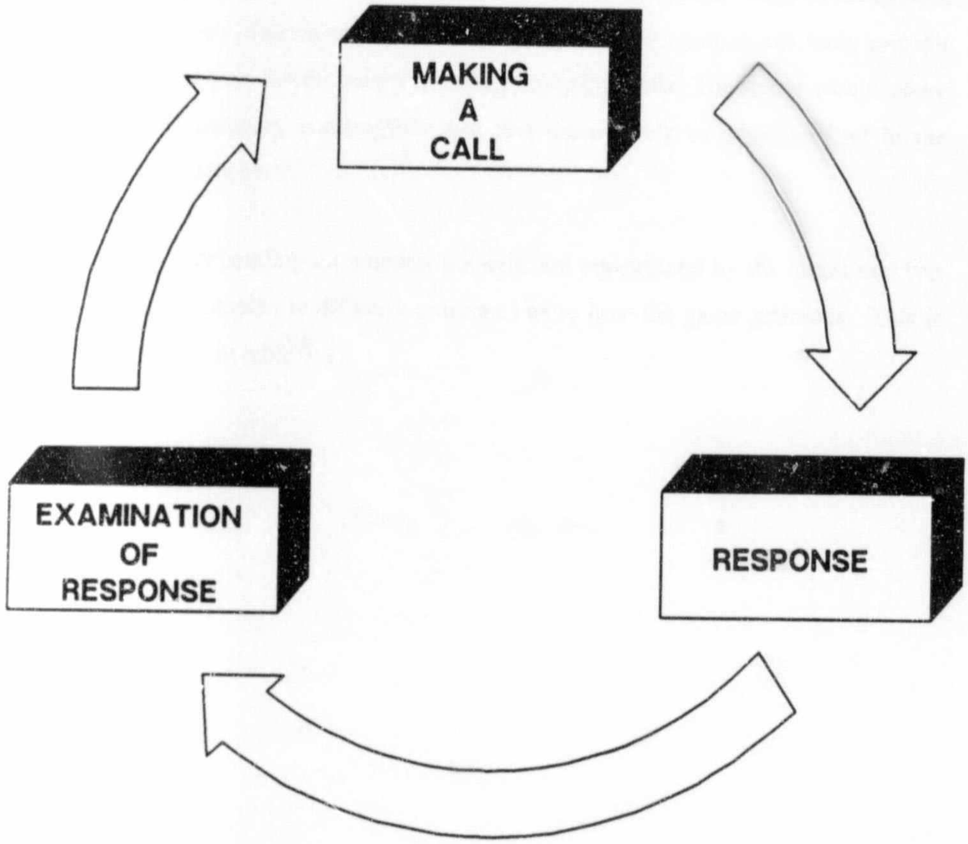

Figure 6-2. Negative feedback in Target.

Let us contrast this negative feedback behaviour to what would happen if there was no connection between the target number and the calls made. We would expect no feedback, and hence no reduction in the range of acceptable calls. Even after several calls had been made, we would not have enough information to decide between widely differing calls. The game would show little exclusionary convergence and this corresponds to what we find in the pseudo-sciences.

So, depending on whether the calls are constrained by the target number, there are completely different outcomes as to how the game proceeds. This is represented in table 6-1.

Table 6-1. Expected patterns in the history of "science" and "pseudo-science" in target.

Discipline	Premiss	Expected pattern
"Genuine science"	Calls are constrained by the target number.	Diminishing call diversity.
"Pseudo-science"	Calls are not constrained by by the target number.	Stasis in call diversity.

Discussion: role of target in demarcation

In chapter 1 a *historical* unit of appraisal was defended. It was argued that there are advantages in departing from the traditional approach to demarcation by seeking an indicator of demarcation in the outcome of method, and so by-passing the question of method. Target helps us to see that the problems of demarcation (in this new sense) and method, though related, are distinct and that there are advantages (and limitations) in by-passing method when devising an indicator of demarcation.

Method must be distinguished from demarcation because a variety of methods may be successfully used in science. One might define success as 'approaching the truth.' For our purposes we can define success in terms of exclusionary convergence. It is easy to see in target that several methods can lead to success in this sense. A method does not have to be the best method to be successful. Naive falsificationism, because it eliminates, would lead to exclusionary convergence. Even the following trial and error method would work: A player could reject all cases except that one that gives a response of four bulls i.e. that is the target number. Of course, this all or nothing approach would be silly. It would take a long time, up to 5040 calls, but the target number would be reached, and the game would show exclusionary convergence albeit rather weakly.

Basing an indicator of demarcation on the pattern that the game as a whole takes avoids the complications of taking into account the great diversity in what may be acceptable method. The same may be true of science.

Target also helps us to see that certain patterns in science may not be easily detectable. Our indicator requires us to have a birds eye view of science. This is difficult because we have to avoid perceiving the past through contemporary spectacles.

If Target is played in a properly "scientific" manner one should expect convergence in the sense of a reduction in the range of acceptable theories over time. This convergence is easily seen because the target number is small and it takes relatively few steps to discover it. If the target number was extremely large, and hence more closely analogous to science, this convergence would be less readily perceivable, particularly if only a few steps of the game were followed. Now for many philosophers(e.g. Stove 1982:3-5) the phenomenon of convergence to the truth in science is an intuitively obvious fact. Yet some philosophers (e.g. Kuhn 1970:170) have even denied that it occurs at all. If there is such a thing as convergence to the truth, Target suggests why this may not be easily detectable. To detect exclusionary convergence (and perhaps convergence to the truth) one is required to examine the history of a discipline over a sufficiently long period of time.

Our analogy may also give us insight into the disputes over what demarcation criteria to use. One of these is the level of agreement or disagreement in a field.

Target and 'level of agreement' as an indicator of demarcation

Creationists often interpret the disputes among evolutionary biologists as signalling the demise of the theory (e.g. Anonymous 1985:14-24). Evolutionists (e.g. Godfrey 1981, 1983) reply that the disagreements between evolutionists are relatively minor compared to what they agree on. Target, as we shall see, also exposes the fallacy in the creationist argument. It reveals that agreement or disagreement *per se* does not indicate much about the credentials of a discipline. The conditions of both science and pseudo-science can be simulated, and we can determine whether 'level of agreement' differentiates between them.

Consider the following two simulations of pseudo-science:

Simulation A: A group of people play target and have to agree on a call before it is made. Further, let us that they do not take much notice of the responses made to their calls. If one asked them individually what they thought was the target number we would expect substantial disagreement between them.

Simulation B: This simulation is as before except that all members of the group have a prior commitment that a particular digit is in the target number. In contrast to the previous situation we would expect substantial agreement between them as to the identity of the target number.

Now we would not want to say that either group were practicing "science". Yet the level of agreement is very different in each case. The level of agreement in a discipline, then, is not a good indicator of its scientific status. What the two situations have in common, and this is what effectively identifies them as pseudo-sciences, is that in neither case do we find a diminishing of call diversity.

Consider the following simulation of science:

Simulation C: As before the group have to agree on a call before they can make it. This time however, they let the responses speak to them concerning the appropriateness of their calls. If we asked them what they thought was the target number after 2 calls had been made (say) we would expect a certain amount of disagreement. If we asked them again after 4 calls had been made we would again expect disagreement, but it would be less in extent. The overall result, in contrast to the previous two simulations, will be diminishing call diversity over time.

Target, then, seems to be a useful tool in exposing the superficiality of the argument that because there are disagreements between evolutionists, the end of evolution is nigh. At the same time it reveals that a history of disagreement over fundamental issues is significant. That it occurs in creation science hints at its pseudo-scientific nature.

In summary, Target is a helpful analogy with regard to the problem of demarcation. It illustrates the distinction between demarcation (in the sense of correctly distinguishing science from pseudo-science) and method; it illustrates the advantages and limitations of by-passing method in devising an indicator of demarcation; and can be used to evaluate demarcation criteria.

2. Target and methodology

Introduction

In chapter 2 it was argued that MCT is distinct from and superior to versions of verificationism and falsificationism because it tells us more. We shall see that in the game it is possible to determine how much more MCT tells us. We cannot, of course, extrapolate these values to scientific practice. What will be evident is that methods in the game can be evaluated by a number of criteria and that these criteria are not equally good. Secondly, we will see that Target illustrates the interdependence between theory and observation. Finally, Target is helpful for illustrating some of the distinctions held by many philosophers of science.

1) Evaluating methodological criteria

The information acquired from each of the three strategies can be determined using probability theory. Given a number of steps and a particular strategy, the information acquired can be taken as the probability of the next call being correct. Information content is inversely proportional to the number of possible permutations or

$I = 1/P$ where I = information

P = number of possible permutations

^nP_r can be determined from the following formula:

$^nP_r = n!/(n-r)!$ n=number of digits taken r at a time

With no feedback, the probability that a given number is the target number is 1/5040 (n=10, r=4) and this is equal to the amount of information we have, I. I≠0 because the target number is not infinite.

The information acquired from the verificationist strategy.

Let us enter the game after two calls have been made. The target number, which is undisclosed to the caller, is 7605.

target number	7 6 0 5	
call #1	0 1 2 3	1 hit
call #2	6 1 2 3	1 hit

The verificationist might reason that if 2 is contained in the target number, then a call of 6023 should elicit a response of at least one hit.

target number	7 6 0 5	
call #1	0 1 2 3	1 hit
call #2	6 1 2 3	1 hit
call #3	6 0 2 3	2 hits

This is in fact what happens. Therefore digit 2 may be in the target number. This is a valid argument but how much has been learnt? Let us find out by applying probability theory. The verificationist does not know which of the digits are the two hits. The number of possible combinations of two digits that could be the hits can be found from the following formula:

$$\binom{n}{r} = r!/(h!(r-h)!)$$

where h = number of hits.

The number of possible combinations for the remaining two positions given the remaining (n-r) digits is :

$$\binom{n}{r\text{--}rh} = (n-r)!/((r-h)!x(n-2r+h)!)$$

The total number of possible permutations knowing which digits are in the target number, but not their correct positions is r!. In this situation he does know something about the positions of the two hits: that there is only three possible positions they could occupy. This limits the possible permutations. A certain value 'y' needs to be subtracted from r!.

The information acquired from the verificationist strategy (V) can be found from

$$1/V = \binom{r}{h} \times \binom{n}{r\text{--}rh} \times (r!-y) \qquad (1)$$

The information acquired from the falsificationist strategy.

Consider, now, how a falsificationist might proceed given the following calls:

target number 7 6 0 5

call #1 0 1 2 3 1 hit

call #2 6 1 2 3 1 hit

He might reason that if 1 is contained in the target number then a call of 6023 should give a response of no hits.

call #3 6 0 2 3 2 hits

He does not get a response of no hits, therefore 1 is not in the target number. Again, we have a valid argument. How much has he learnt? Since he has excluded one digit from further consideration we must substitute

$\binom{n-r-1}{r-h}$ for $\binom{n-r}{r-h}$ in (1)

so that the information acquired from the falsificationist strategy (F) can be found from

$$1/F = \binom{r}{h} \times \binom{n-r-1}{r-h} \times (r!-y)$$

where

$$\binom{n-r-1}{r-h} = (n-r-1)!/((r-h)!x(n-2r+h-1)!)$$

The information acquired from MCT.

Consider how a follower of MCT might proceed given the following calls:

target number 7 6 0 5

call # 1 0 1 2 3 1 hit

call # 2 6 1 2 3 1 hit

MCT requires that one imagine alternative explanations that will account for the data. After the response to call #2 there are two possibilities.

A_1 : 0 and 6 are hits (i.e. contained in the target number) and 1, 2 and 3 are misses (not contained in the target number).

A_2 : either 1, 2 or 3 is a hit and 0 and 6 are misses.

These alternatives represent scientific hypotheses. We need to discover which is correct. This can be done by devising a call for which the two explanations predict different responses. I will call such a call a discriminatory call. A discriminatory call for explanations A_1 and A_2 would be 6 0 2 3. For this call one would expect that if A_1 is correct then the response would be 2 hits (or 1 bull and 1 hit). If A_2 were correct then the response would be 1 hit (or no hits).

call # 3 6 0 2 3 2 hits

The discriminatory call (call #3) gives a response of 2 hits therefore explanation A_1 is the correct one. We can note in passing that that call #3 is identical for the three strategies. This illustrates nicely the point made in chapter 3 that having merely the right data may not be enough. The issue may be, as it is in this case, of correctly interpreting the existing data. Furthermore, we see that in order to effectively adjudicate between the hypotheses it is necessary to define these clearly and to follow carefully their consequences. The same applies to scientific research as we saw in chapter 2.

Since explanation A_1 is correct, we can eliminate

$\binom{r}{n}$ in (1)

so that the information acquired from MCT (1/M) can be found from

$1/M = (^{n}{}_{r\text{--}rh}) \times (r!\text{-}y)$

The ratios of the amounts of information acquired through the 3 strategies can be found since n, r, and h are known and (r!-y) can be eliminated as it is constant for all methods.

substituting we get:

$F/V = 1.5$

and

$M/V = 6$

Now, the reasoning behind each strategy is equally valid. Yet, according to the amount of information received from each method the falsificationist strategy is 1.5 times and the multi- cornered strategy 6 times better than the verificationist strategy given the above values of 'n', 'r' and 'h'. This was possible because of the way the game was set up. These values of 'n', 'r' and 'h' were predetermined. In scientific practice, as Ghiselin (1974:4) has pointed out, we do not know if we have considered all possibilities. We cannot determine 'n' in scientific practice. The ratios obtained for the amount of information obtained from the various strategies cannot be extrapolated, of course, to their corresponding methods in science since they depend on the values of 'n,' 'r' and 'h'. What is significant is that despite the underlying logic being valid in all cases, they were not equally effective. The most effective strategy seems to be one in which two or more "hypotheses" are forced to compete. Thus, the game illuminates the point that methodological criteria such as 'the underlying argument is valid' are limited in their usefulness. The same can be said of the criterion 'it works' for the following reason.

We have seen that even a comparatively bad method, the trial and error method, is successful. This suggests one should not make a method normative just because it is successful. Several authors (e.g. Lakatos 1970:177-80, Stove 1982-4) have accused Kuhn of doing just this. However, Kuhn (1970: 236-237) has said that a method should be taken as prescriptive not just because it is practiced and works, but because it fits a theory of science. However, accepting this criterion along with others may lead one astray. We have seen that the question to ask of a method is whether it is the most effective option.

Target can also be used to illustrate the interdependence between observation and theory.

2) Illustrating the interdependence between observation and theory

Consider the following steps:

target No. 7 6 0 5

call # 1 0 1 2 3 1 hit

call # 2 6 1 2 3 1 hit

call # 3 6 0 2 3 2 hits

call # 4 6 0 4 3 2 hits

call # 5 6 0 5 4 3 hits.

Let the calls represent theories, the response represent observations and the target number represent the real world. The response is only made after the call. Theory precedes observation. But with the exception of the first call, the calls are not wild guesses. You could question the good sense of the calls but they are not entirely inappropriate. Each call barring the first was made in the light of the responses obtained from previous calls. Similarly theory is constrained by past observations and theory. There is a sense in which observation precedes theory. The relationship between call and response is similar to the relationship between theory and observation; it is one of interdependency.

The response to each call does not only depend on the call but also on the target number. There is a sense, therefore, in which the response is independent of the call made. Similarly, observation is partly theory independent. The relationship between the calls, responses and the target number are depicted in figure 6-3.

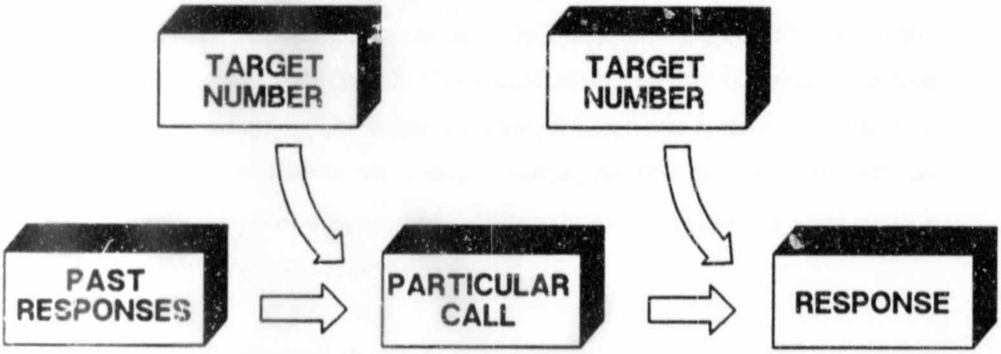

Figure 6-3. The relationship between calls, responses and the target number.

Analogies are often helpful for illustrating some of the distinctions that have been suggested by philosophers. Below we will consider a few of these distinctions:

3)Illuminating distinctions using Target

One distinction is between the context of discovery and the context of justification (Hanson 1968, Popper 1972a:31) The idea here is that no matter how a hypothesis is originally arrived at, the testing of the hypothesis is a separate matter. The parallel in evolution is that the origin of a variant is unrelated to whether it will be adapted or not (Toulmin 1970:46). There is a similar parallel with Target.

One can imagine a player choosing his year of birth for his first call. By sheer fluke he may be right, but he can only justify believing this if his opponents response accords with his call. The game further suggests that this distinction does not mean it does not matter at all how a hypothesis is arrived at. While an element of luck may favour a player, barring the first call, he must still be constrained by previous responses, for him to be likely to reach the target number in an acceptably low number of steps.

Philosophers of science (e.g. Newton-Smith 1981:4-8) have made a distinction between the reasons scientists choose particular hypotheses, and the reasons they *should* choose hypotheses. These reasons may not always coincide. One can easily imagine a situation where a player uses the fallacy of affirming the consequent to conclude that some digit is in target number, and he happens to be right. This does not mean that his reasoning was correct. The game can be used to distinguish between the correctness of a conclusion and the soundness of the argument.

There has been a long debate over the cognitive status of theories (Nagel 1961: 117-152, Barbour 1966:162-171, Popper 1976: 97-119). The major views are the instrumentalist view in which theories are merely tools for making predictions; the idealist view in which theories are human constructs; and the realist view in which theories are human representations of the real world. Depending on how closely the Target analogy fits science, it seems to support the realist view that theories are multifarious.

Each call is the invention of the player. Similarly the idealist is right when he says that theories are human inventions. However, he is wrong to make this a complete description of the theory. A call may or may not agree with the target number. The realist has a right to ask whether a theory represents reality or not.

Gardner (1983) argues that the role of a theory can change. He says that Copernicus' theory was initially accepted as an untrue prediction device, but later came to be regarded as true or approximately so. In a similar way, the role of a call changes as the game proceeds. At the beginning, the call is most evidently a useful fiction. It is used as a tool for making predictions- the player reasons (for instance) that if call x is made and digit y is not in the target number, then the response will be z. He does not seriously believe that x exactly matches the target number. Towards the end of the game, as the evidence accumulates, the call is presented more and more as a claim to match the target number.

CHAPTER 7: CONCLUSIONS

The problem of demarcation

The first major objective of this thesis was to develop an indicator of demarcation which met the following conditions: It must meet the urgent need for correctly identifying genuine sciences and pseudo-sciences, but at the same time must be free of the main problems involved in demarcating on methodological grounds. It need only be applicable to disciplines which claim to be genuine sciences: it was not designed to deal with non- sciences. It must successfully determine the status of evolutionary biology and creation science. In this chapter the merits and limitations of a historical indicator of demarcation are discussed by determining how it, as opposed to other criteria, solves certain problems. We shall find that along with its limitations it has certain special advantages. Later in the chapter scientific method and the practices of evolutionary biologists will be discussed.

One problem with demarcating on methodological grounds is that philosophers of science are not in complete agreement over what scientific method is. Some dispute whether there is a 'the scientific method' (e.g. Feyerabend 1975). It seems that whenever a method is suggested as the hallmark of science, it is later discovered that some pseudo-science or non-science practices it, while some apparently genuine science does not.

Consider the notion of revisability. This was supposed to be a feature of the genuine sciences alone. Yet Laudan (1982:151) has noted that the creationists have revised their position regarding the amount of variability at the level of species change. Furthermore, revisability is not always present in the genuine sciences. In the early development of Newton's gravitational theory it was discovered that the theory did not accord with observations of the moon's orbit. Newtonians tried to save the theory in the hope that the problem lay with something other than its core tenets. This hope was satisfied.

Or consider Kuhn's (1970:8,9) proposal that science is characterized by a puzzle-solving activity. Feyerabend (1970:200) criticized this criterion on the grounds that it could not exclude organized crime and Oxford philosophy.

With our indicator we can correctly distinguish genuine sciences from pseudo-sciences, while respecting the lack of consensus about method, and without waiting for a completely satisfactory characterization of scientific method. Scientific method is placed in a black box with exclusionary convergence as the outcome. This is not to say that method is utterly mysterious or that it should not be used in demarcation. What is suggested is that we can correctly identify a discipline's status without even looking at its methods.

A further problem is that there is some diversity in what is acceptable and unacceptable in method. We have seen that in certain fields such as evolutionary biology emphasis is placed on statistical testing. Deductions in the historical sciences are more likely to be retrodictions than predictions, and a study of background conditions assume a greater importance in biology and geology than in the 'hard' sciences.

There is diversity in the bad methods of the pseudo sciences. One tactic used by the creationists is appeal to miracle (Gould 1984:127). Popper (1972b:37-38) says that astrologers make the mistake of always looking for confirming instances. It is also unacceptable to retain falsifiable but false hypotheses.

One way to deal with this complexity is to specify a set of conditions, only some of which need be present to classify an enterprise.

One merit of our historical indicator is its simplicity. It is effective despite this diversity. For, whatever the legitimate methods of a science are, we can expect one outcome- exclusionary convergence, and whatever the illegitimate methods of a pseudo- science are, we can again expect one outcome- little or no exclusionary convergence.

Another problem is that pseudo-sciences may be such for a variety of reasons. It was mentioned earlier that some parts of creation science are falsifiable and false. The claim that earth and the rest of the universe is about 10 000 years old is false. Other parts, such as the claim that the earth was created with apparent age, is unfalsifiable in principle. This suggests that different parts of creation science must be rejected for different reasons.

A further merit of our indicator is that it avoids these complications. The discipline as a whole will show no exclusionary convergence in spite of this disparity. For the parts that are false we expect no exclusionary convergence, because the pseudo-scientists refuse to allow changes to their theories. For the parts that are unfalsifiable we again expect no exclusionary convergence, but for a different reason. There is no empirical evidence which would force them to change their theories even if they wanted to.

There is a further problem with demarcating on methodological grounds. Pseudo-scientists can, as is well known, pay lip-service to method, and superficially appear to be scientific. Creationists have had some success with this method.

Raup (1983:159) says: "A few of the arguments used by the creationists "are scientific" in the sense that they use the basic methods of testing hypotheses normally considered to be scientific. This does not mean, of course, that the conclusions are correct. Bad science may be difficult to distinguish methodologically from good science."

Creationists often appear to argue along the line of multi- cornered testing. The 'predictions' of evolution and the 'creation model' are laid side by side. Table 7-1 is part of one set up by a creationist (Anon. 1985:55):

Table 7-1. Some "predictions" from evolution and scientific creationism

"Model"	Predictions
"Orthodox evolutionary theory anticipated a fossil record that contains:	"... Very simple life forms gradually appearing ..." "... Many transitional "links" between different kinds..."
"...The creation pattern anticipated a fossil record that contains:	"... Complex life forms suddenly appearing... No transitional links between different biological families..."

Kitcher (1982:38-42) gives one response to this mimicing of the sciences by the pseudo-sciences. He has thoroughly examined creationist method, showing that the resemblance to genuine method is only in appearance. For instance, some of their predictions are not derived from evolutionary theory alone, but from this theory with other faulty assumptions. Kitcher rejects creation science not because it does not satisfy some hallmark of science, but for a variety of methodological reasons requiring a detailed examination of their methods.

Our indicator solves the problem of lip-service in a different way. Since it is only an indicator, it cannot actually identify illegitimate methods. However, this limitation is offset by one useful advantage. It is not susceptible to pretenses in scientific method because it judges a discipline by its fruits. A discipline will not show exclusionary convergence whatever the tactics used to cover up its ignoring of the evidence. Since it is a historical indicator, whatever creationists do with method, they cannot alter the history of their discipline.

Finally, there is the problem that progress in science has not always been smooth. Aristarchus' heliocentric model of the solar system was rejected in favour of the Ptolemy's geocentric model. While there were anticipations of evolutionary theory before Darwin, these were swamped by the rise of natural theology and independent creation in the seventeenth and eighteenth centuries. In chapters 4 and 5 we saw that evolutionists, in this century, sometimes rely on less than the best in method.

There are several possible responses to this problem. We could admit that evolutionary biology is not in the same league as the physical sciences. However, this would give creation scientists additional ammunition against evolution.

A historical indicator gives an alternative solution. It allows us to reject poor methods, but at the same time accept evolutionary biology's scientific status for two reasons. Firstly, the unit of appraisal is the discipline as a whole and not parts of it at particular moments. We can make allowances for mistakes as long as these do not keep the discipline as a whole from displaying exclusionary convergence. Secondly, demarcation and method, although related, are treated as distinct issues. This is because less than the best in method can still lead to success- to exclusionary convergence.

Naive falsificationism would lead to success because theories would still be eliminated. Yet, naive falsificationism is not a good method. Method and demarcation, should, then be distinguished. To claim that a discipline is a science is to say that it has made the qualifying time for entering the scientific race. How it fares in the race depends on how good its methods are. Once a discipline is found to be a science, a second higher level criterion must be applied to its methods- are these the best available for acquiring information.

The fact that even apparently bad methods can lead to convergence suggests that there is no such thing as *the* scientific method. Bad science is still science. To claim that a discipline is a science would, then, be a weak claim. It would be merely asserting that the discipline has made the qualifying time. It has entered the scientific race. How it performs in the race depends on how good its methods are.

Because a particular method works does not mean we should assign it a prescriptive role. Another higher level criterion must be applied to it - is it the best method out of the alternatives available; best at acquiring information?

The problem of whether some higher level criterion could be developed for method was therefore treated separately and will be discussed in the next section.

The problem of method

Philosophers of science today evaluate methodological guidelines according to two sets of criteria. Firstly, whether they can be given an *a priori* justification- whether they can be justified in principle. Secondly, whether they can be given an *a posteriori* justification- whether successful research is based on these methods. The objective of the next section is to evaluate MCT according to both sets of criteria. We will find that there are good *a priori* reasons for accepting MCT. Furthermore, we will find that Darwin's research was so successful partly because it approximated MCT, and creation science is so unsuccessful partly because it does not. At the same time possible objections to this method will be discussed. Next, the methods used by evolutionary biologists will be discussed. We will find that there are reasons why relying on common evidence is so seductive, and consequently occurs quite often, and why the significance of discriminatory evidence is not always apparent.

***A priori* considerations**

The purpose of chapter 2 was not to justify MCT as some new reliable methodological algorithm. MCT is not an infallible method. I have merely argued for its distinctiveness and superiority over two other procedures that biologists use.

The evaluation of the three procedures was done *within* a deductive perspective. It should not be seen as an attempt to lobby for the superiority of deductive over inductive procedure. Philosophers of science are not agreed over whether these are distinct methods or different aspects of the same basic method. Neither have I attempted to discuss all of the many methodological principles that have been suggested by philosophers of science. The focus of this thesis has been empirical rather than rational criteria in theory choice. Philosophers of science agree that the latter criteria, such as simplicity, coherence and explanatory power, are important (Hempel 1966, Koertge 1978, Kitcher 1981, 1985). However, it is not always clear which criteria are used when scientific research is successful (Lloyd 1983). Rather than become embroiled in this controversy my object has been to emphasize the benefits of theory comparison using empirical criteria as an example.

It is important to dwell on the distinctiveness and superiority of theory comparison because, as we saw in chapter 5, there are cases where this approach is conflated with others. Furthermore, it seems that philosophers of science have not given this method its due in contributing to the success stories in Evolutionary Biology. In the context of Darwinian histories, for example, Kitcher (1985:66) says that philosophers of science have not sufficiently emphasized that these "win their way to the top through the elimination of rivals."

In chapter 3 it was argued that the validity/invalidity criterion is not the only or the most useful in evaluating method. According to Stove (1982) a preoccupation with logical validity leads to a perfectionism and skepticism which is actually a hindrance to science. I suggested that we should rather ask how much a method can tell us. From the perspective of this criterion, MCT was shown to be preferable and distinct.

It was noted that certain complications in the actual practice of science forces a restructuring of the basic MCT argument. This modified argument seems useful because it provides a formal yet simple framework for understanding an existing method. At the same time it avoids certain pitfalls. It emphasizes that we can never say that a theory is verified, but we can ensure that our theories approach the truth; it is important to consider alternative explanations for a given set of facts; it is important to subject these explanations to a series of multi-cornered tests.

To say that MCT is the preferable method of the three is not to say that it is the only valid way to proceed. It was pointed out that the verificationist argument can be rephrased to convert it into a valid one. The mistake is to make out that the results of a particular method are more significant then they actually are. To put it differently: To say that MCT is the preferable procedure is not to say that collecting common evidence is of no value, but rather that it is of limited value. What is unacceptable is to make out that it has more significance than it actually has.

The distinctive feature of MCT as compared to naive verificationism or falsificationism is that it is comparative. The relative merits of alternative hypotheses or theories are compared. This leads to the consequence that there are two distinct kinds of favourable evidence- common and discriminatory.

Two objections have been levelled against the comparative approach to method. Answering these objections will help to clarify the role of this approach. Lewin (1983) quotes May as saying that "it is paradoxical that some of those who are most sensitively aware of the need to keep sight of alternative explanations ... seem, at the same time, occasionally to accept that there is only one True Way to do science." Others might say that this approach assumes that out of the available hypotheses only one is true. Both criticisms would be based on a misunderstanding of the role of the comparative approach. MCT is not an empirical proposition- it does not commit the investigator to believe that it is single causes or many factors that are behind particular biological phenomena. MCT is a investigating procedure that is heuristically advantageous- by considering alternatives we are more likely to discover what is the case.

The advantages of the comparative approach

MCT fosters a fallible yet ambitious approach to science. It suggests that while we may not be able to have any absolutely certain knowledge, we can know that some theories are better than others. We will maximize progress by forcing theories to compete. MCT is at once both modest and critical. Modest because the claim is not that the method produces proven theories. Critical because it does not accept just any theory. A theory must prove its worth, not against some unreal standard but against the current best theories. So demands are made on theories but these demands are realistic. To see the benefits of this approach let us compare it with the alternatives.

At the back of naive verificationism is the view that the only significant outcome in science is proof. This has some unhappy consequences that have become very apparent when viewed against the back-drop of the creationism-evolution debate. Kitcher (1982:31-35) and Kehoe (1983) show that creationists have had some success in their lobbying for equal time because of their strategy of relying on the popular misconception that scientific knowledge is absolutely proven knowledge. Then they point out that evolutionary theory has not been proved and therefore cannot count as a genuinely scientific theory. Philosophers of science have repeatedly pointed out that if these stringent requirements are strictly kept, virtually all of generally accepted scientific knowledge would have to be thrown out.

The antidote for such gratuitous demands for what can count as a science is a comparative approach to method. It needs to be emphasized that for it to be reasonable to accept a theory one has only to show that it is the best available; one does not have to prove its truth in an absolute sense.

At the back of naive falsificationism is the view that the really significant outcome in science is the elimination of error. The elimination of error is an important task of science, but a pre-occupation with this task has unfortunate consequences. Creationists have made some headway by focusing on the errors and deficiencies in evolutionary theory. The comparativist would be immune to this tactic. He is able to admit problems with the theory without feeling compelled to reject the theory.

A Posteriori considerations

In the next section it will be argued that Darwin's method was successful partly because his method approaches that advocated in chapter 3, whereas creation science is unsuccessful because their method does not. In so doing, my intention is to provide additional *a posteriori* justification for the comparative approach to method.

Darwin's method

There have been several interpretations of the structure of Darwin's method in the *Origin of species*. These include hypothetico-deductivism (Ghiselin 1969, Ruse 1973:210, 1982:53- 57), inference to the best explanation or concilience of inductions (Thagard 1978), the semantic view (Lloyd 1983) and the causal efficacy view (Recker 1987).

It is not necessary for my central thesis that I take sides here. My one object is to expose the inadequacy of the creationists approach to method. For this, all I need to show is that Darwin's argument has a comparative structure, whereas theirs, in reality, does not.

With all their differences, there is one point on which all the above authors agree. All accept that Darwin argued for the superiority of his theory over its rivals. For Thagard (1978), Darwin saw his theory as the best explanation of the available facts. For Lloyd (1983), Darwin saw his theory as the most empirically adequate. For Recker (1987), Darwin saw his theory as having the most explanatory power and as supplying the *vera causa* of life.

In the ensuing discussion I will be placing Darwin's argument in deductive clothes. This is purely for convenience. I do not mean to champion hypothetico-deductivism over the other interpretations of Darwin's method. Evidence that Darwin's method had a comparative structure comes from a number of lines.

Firstly, we see this in his attitude to anomalies. One would expect the naive falsificationist, in contrast to the comparativist, to exaggerate the force of the anomaly. This is because he is only concerned with how a single particular theory squares against the real world. Instead, Darwin writes with disapproval that "Any one whose disposition leads him to attach more weight to unexplained difficulties than to the explanation of a certain number of facts will certainly reject my theory."

Secondly, Darwin tacitly recognized a distinction between common and discriminatory evidence. There were a number of facts about the biological world, readily predicted by his theory but which he did not use in its defense, because it was obvious that a rival predicted these just as well:

i) Sterility versus degrees of sterility.

Sterility is something which is readily explained by Darwin's theory as an incidental by-product of other biological processes. Darwin did not invoke sterility as such in the defence of his theory. He knew that it could be deduced from both theories. He wrote (1859:264): "The view generally held by naturalists is that species when intercrossed, have been *specially endowed* with the quality of sterility, in order to prevent the confusion of organic forms."

Instead, Darwin appealed to the fact that sterility between organisms *is in degrees*. He said (1859:436): "We see the truth of this conclusion [that sterility is incidental on constitutional differences rather than a special endowment] in the vast difference in the result, when the same two species are crossed reciprocally

..." Under Independent Creation one would expect absolute sterility between species, irrespective of the direction of the cross.

ii) Close fit between organism and environment versus imperfections.

He was well aware that his theory could account for the marvellous design in living things but he did not appeal to design as such for support of his theory. We do not see Darwin constantly remarking on the fit between organisms and environment. This must have been because he knew that such evidence, as it stood, could be accounted for by Independent Creation. Rather we find him focusing on the lack of fit between organisms and environment, e.g. sub-aquatic thrushes and upland geese (1859:216-7). It was *this* evidence that was discriminatory.

iii) Diversity versus unequal distribution of diversity.

Diversity as such is something readily explained both by Darwin's theory and by Independent Creation. According to the former diversity is a consequence of adaptation to different environments. According to the latter variety was good in itself and reflected the abundance of Providence (Ghiselin 1969:44). Again we do not find Darwin appealing to diversity as such. Rather, he points to the way in which this diversity is distributed. With Darwin's theory one would expect organisms to diverge with increasing geographical separation. Under Independent Creation one would expect organisms to diverge with increasing divergence in conditions. Darwin, thus stresses the fact that "neither the similarity nor the dissimilarities of the inhabitants of the various regions can be accounted for by their climatic and other physical conditions" (1859:344).

Three instances are given above but Darwin consistently thought this way. He repeatedly says things like: "how inexplicable on the theory of creation is the occasional appearance of stripes on the the shoulder and legs of the several species of horse-genus and in their hybrids! How simply is this fact explained if we believe that these species have descended from a striped progenitor..." (1859:446).

Not that Darwin thought that there existed a single or a few lines of crucial evidence that would conclusively verify his theory over rivals. In his section on instinct in the Origin of Species Darwin (1859:262,263) says that he does "not pretend that the facts in this chapter strengthen in any great degree my theory ..." but that "... all tend to corroborate the theory of natural selection."

There must have been some realization here that it was rather the collective evidence that tended to favour his theory over rivals. This is so even though Darwin did not argue this neatly, or, as Hull (1973:25-26) has shown, did not consistently appreciate that each line of evidence does increase the support of a theory without absolutely proving it. Figure 7-1 places Darwin's reasoning into a deductive framework:

If Descent by Modification is true. then

B_{11}: Greater variation in species rich groups.

B_{12}: Imperfectly adapted animals. eg. upland geese with webbed feet, sub-aquatic thrush.

B_{13}: Organisms distributed according to their taxonomy.

B_{14}: Existence of rudimentary organs

etc.

If Independent Creation is true. then

B_{21}: equal variation in species rich and species poor groups.

B_{22}: Perfectly adapted animals.

B_{23}: Organisms distributed according to conditions alone.

B_{24}: No superfluous organs.

etc.

B_{11}, B_{12}, B_{13}, B_{14} etc.

Therefore Descent by Modification is true (or better).

Figure 7-1. The comparative structure to Darwin's argument.

Scientific creationist's method

In chapter one we saw that there are several types of creationism. One concept is purely theological. The existence of the universe is said to be contingent upon the existence of God. No attempt is made to justify belief in a Creator on scientific grounds. I will not be concerned at all with creationism of this sort. Unfortunately, the same term is used for a quite different concept. In the second category, an attempt *is* made to defend the idea of creation on scientific grounds. This category can be further subdivided. There is progressive creation, in which it is claimed that the fossil record supports the notion of successive creations over the geological time scale; there is scientific creationism in which it is claimed that the fossil evidence supports a recent creation, and there are several other brands. In the following discussion my concern is only with scientific creationism which will, following current usage, also be called creation science or creationism.

In spite of the overwhelming evidence against it, scientific creationism is a growing movement (Numbers 1982, 1987). Numerous authors (e.g. Kitcher 1982, Futuyma 1982, Godfrey 1983, Montagu 1984) have argued that scientific creationists use questionable tactics on their uninitiated audience to achieve this success. They have been accused of distorting the evidence, of misrepresenting evolutionists and of ignorance of both the facts of evolutionary biology and of philosophy of science. The object in the discussion below is not to give an exhaustive assessment of scientific creationism, but to show how selected methodological problems of the movement are the result of violating the methodological principles established in chapter three.

The fallacies in the thinking of creationists are often obscured because the arguments are couched in technical language. That their arguments are meretricious will become more apparent if we leave the controversy for a moment and consider an illustration. We will see that the creationist argument hangs on a confusion between three concepts of evidence. In figure 7-2 let the dots represent facts and the curves hypotheses.

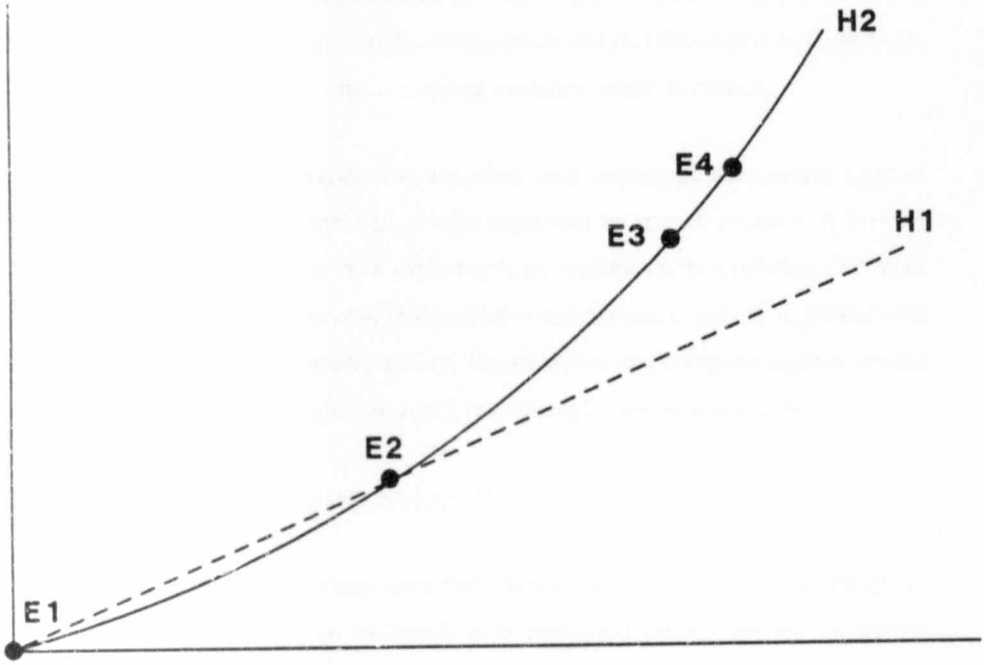

Figure 7-2. Concepts of evidence.

In this case the evidence supports H_2 over H_1. How could someone create the impression that H_1 is better supported than H_2. He might use the following tactics, which, I will show later is just what the scientific creationists use:

a) He could focus on the evidence common to the two hypotheses (E_1 and E_2) to the exclusion of the evidence that discriminates between them (E_3 and E_4). This one type of special pleading. The fallacy is immediately evident in the illustration. It is, of course, of no consequence that H_1 (the weaker hypothesis) is supported by E_1-E_2. It is the remaining evidence which is critical.

Scientific creationists are often very impressed that certain lines of evidence (of the common variety) are explained by special creation. It is often incorrectly claimed that such evidence is inconsistent with evolution. We shall see that this evidence, as many authors have pointed out, is only inconsistent with a caricature of evolutionary theory. The evidence that mitigates against special creation is frequently ignored. Just a few examples are shown below.

i) Apparent design versus imperfections.

We have already seen that Darwin did not use apparent design as evidence for evolution, although it is predicted by his theory. Scientific creationists frequently do (e.g. Clark 1961:141-191, Coppedge 1973:215, anonymous 1986). Coppedge (1973:215) writes: "There is no way chance mutations could have brought about the development of feathers. Each is a masterpiece of engineering." According to evolutionary theory it is not chance mutations that bring about development, but natural selection (a powerful anti-chance force) acting upon variability. Even this is a very simplified picture. That apparent design can be accounted for by evolutionary theory is very well attested to. The dark form of the moth *Biston betularia* is camouflaged against the dark polluted bark of trees. The "camouflage design" is apparent, and merely the result of selection against the pale form. Therefore, as Cracraft (1983:176) points out: "... creationists should not invoke common design because ... statements about functional design ... are consistent with evolution as well."

While creationists push the common evidence of apparent design, they are revealingly silent about the imperfections in living things such as the webbed feet of upland geese, and the absence of webbed feet in sub-aquatic thrushes. Coppedge (1973), for instance, completely ignores the imperfections in living things. It is this sort of evidence which is very difficult to explain on the basis of special creation.

ii) Similarity versus the pattern of similarity in space and time.

Similarity is frequently used to support special creation (e.g. Wilder-Smith 1974:52-53, Frair and Davis 1983:25). Wilder- Smith writes: "Darwinism postulates that similarities in living organisms predicate common ancestry." He suggests that the observed resemblances between living organisms might "be better explained on the basis of a common plan or planner behind them." He does not say why it should be *better* explained this way.

The impression created is that evolutionists are so desperate for evidence that they have to rest on mere similarity as evidence for evolution. Evolutionists need not, and do not rely on similarity *per se*. Darwin appealed to certain similarities between organisms geographically associated but in widely differing conditions. It is *this* evidence that favours evolution over special creation, and Wilder-Smith completely ignores it. For, one would expect an intelligent designer to plan organisms according to their environmental conditions alone. It would be capricious to alter the plan with increasing geographical separation. Further evidence that discriminates is that the similarity in organisms forms a hierarchy. Cracraft (1983:172) argues that "the hierarchical pattern produced by the shared similarities observed among organisms is predicted by a hypothesis of evolutionary descent with modification but not by