

Others were more rigorous in seeking causes for change within the natural order. For example, Empedocles (495-435B.C.) suggested that the parts of individuals united purely fortuitously, with those combinations that were unnatural becoming extinct (Osborn 1929:52-53). Those that accepted a naturalistic agent of change disagreed over causes. Anaximander (611-547 B.C.) thought that the first men were derived from aquatic animals by a process that was roughly similar to natural selection (Gardner 1972:28-29). In contrast, Aristotle "advanced his own philosophy of purposive natural causation" (Osborn 1929:74). Henry of Hesse, in the fourteenth century, wrote that the stars had the power to form new species from old (Clark 1966:33). Leibnitz (1646-1716) attributed species changes to great changes of habitat (Osborn 1929:143). Although special creation came to dominate both science and theology up to the time of Darwin, this should not be taken to mean that there was consensus. A philosophy of change was accepted by many (Osborn 1929) such as de Maupertuis (1698-1759), Erasmus Darwin (1731-1802), Lamarck (1744-1829) and Robert Chambers (1802- 1871).

There is a parallel here with the initial stages in the development of physical optics. There is great diversity in thought. Now our indicator predicts, that for scientific disciplines, there will be in time a limiting of the degree of difference between theories. This is what we find for evolutionary biology. By the time of Darwin many of the older views, such as spontaneous generation and transformism had been discarded. For a time, says Thomson (1895:225) the issue was one which "is now out of date- the question between evolution and non-evolution." Then the scientific community came to reject any supernatural element as an agent of change. More than accepting a merely natural history for life, they came to accept an evolutionary history. Thus Thomson (1895:213) wrote: "Almost all naturalists now accept the Doctrine of Descent...the *fact* of organic evolution forces itself upon us, but a study of the *factors* is still a lesson in uncertainties" (emphasis his). The thinking of the scientific community has converged over two issues at least: stasis versus flux, naturalism versus

supernaturalism. There is still diversity in thought, but it is of a different order. Evolution accepted, the question now turned to its mechanism. The largest school was Darwinian but there were many rivals: a small school was Lamarkian; some proposed unknown laws (e.g. Dawson (1893:341); some would not commit themselves to any mechanism; Gould (1983:143) reports that some actually suggested that the best designed organisms were not the fittest, but *must* perish.

Today there is convergence over what the agents of change are- natural selection, genetic drift etc. The disagreement that exists is of a different order- over the relative importance over these agents of change. There has been exclusionary convergence, then, in *terms* of the agent of change. Let us consider one further aspect- the rate of change.

Linnaeus (1707-1778), as is well known, believed in the immutability of species. Gottfried Wilhelm Leibnitz (1646-1716) presented a doctrine of continuity along which advance is made by small steps (Mayr 1982:325). At the same time many accepted transformationism, (e.g. de Maillet, 1656-1738) the idea that great changes could take place in nature such as fish turning into birds (Osborn 1929:163). Spontaneous generation, the idea that living things could be generated spontaneously from mud or carcasses, was continuously held since the time of the early Greek philosophers (Osborn 1929:18, 48-50).

By the early 1900's, there was still much disagreement over the rate of change. Darwinism upheld a gradual evolution but Nageli (1871-91), Cope (1840-91) and Mivart (1827-1900), amongst others, propagated saltationism- the idea that species evolved in large abrupt steps (Singer 1931:305,306, Hull 1974:51,52). As late as the 1940's, there was still place for Goldschmidt's hopeful monster theory, - that variations in nature were large, although his theory soon met with disfavour. There is much controversy over the rate of change of evolution, but no-one accepted the extremes that were viable centuries earlier.

Even instantaneous creationists accept micro-evolution, that small changes do occur. Biologists rejected a static living world because no matter how the fossil record was interpreted it was clear that change must have occurred. Neither was the rapidly changing world of transformationism and spontaneous generation acceptable at the beginning of this century. In terms of the rate of change in evolution there has been a reduction in the degree of difference between theories.

This trend has continued up until today. There is a controversy over whether evolution proceeds in small jumps or gradually. This controversy can be misunderstood. It is not that Goldschmidt's hopeful monster theory has re-emerged with new hope. As Kitcher (1982:149) writes "there are significant differences between the standard interpretation of Goldschmidt's ideas and the proposals about macromutations that are currently being discussed." While there is controversy over the rate of change in evolution, there is a much narrower range of possibilities acceptable today than there were four decades ago.

In terms of the agents of change and the rate of change the history of evolutionary ideas conforms to what would be expected of a science and parallels the historical development of physics. Our indicator of demarcation reveals that evolutionary biology is a genuine science.

Is creationism a science?

It was noted previously that during the late nineteenth and early twentieth centuries the scientific community came to reject supernatural explanations of the origin of life. These explanations can perhaps be viewed as scientific theories that have been proven false. At this point there is a fork in the road. One road was taken by the orthodox scientific community. Their passage, we saw, was marked by exclusionary convergence at least with respect to the extent of change, and the mechanism of change. The other road was taken by a small group outside of the

orthodox scientific community, which continued to support supernaturalistic accounts of the origin and history of life. Most theologians came to accept Darwinian evolution. These defended creationism on purely theological or metaphysical grounds. I will not be concerned at all with this version of creationism. Instead, I will examine, very briefly, the history of the heterodox group, which attempted to justify creationism on scientific grounds. I will refer to this version as 'creationism.'

The creationist response to Darwin and the findings of geology in the late nineteenth century was very pluralistic. Below is a cross-section of some of the ideas promulgated at the time:

a) Literalism. According to this view, held for instance by G. M. Price (1923) and Gosse (1857) the world was created about 10 000 years ago in six days of 24 hours.

b) Successive catastrophes and creation. Agassiz believed that there were sudden periods of catastrophe, then creation, then a long period of quiet, with this pattern repeating itself again and again.

c) Flood geology. Here the fossil record is interpreted as the remnant of Noah's flood (Price 1923:686-691).

d) The Gap Theory was advocated by Thomas Chalmers, Sedgewick and Pember amongst others (Ramm 1954: 135). Although there were differences of interpretation the general view was that there was a creation, followed by a long period of time, then a catastrophe and finally a recreation.

e) Local creation. J.P. Smith held that the six days of Genesis referred to a divine rehabilitation of just a small portion of the earth (Ramm 1954:131).

f) The Apparent Age view. Gosse simultaneously accepted Literalism and the findings of geologists. He believed that the fossils did not really exist

but were created giving the appearance that the earth was very old (Birx 1984:90,304).

g) Progressive creationism. Evolution is only accepted within the major phyla. Supposed gaps between the phyla in the fossil record is said to support a supernatural explanation for the initial appearance of the phyla.

h) Directed Evolution. Robert Chambers (1890:153-154) said: "I contemplate the whole phenomena (of the development of simple organisms to more complex ones) as being in the first place arranged in the counsels of Divine Wisdom, to take place, not only upon this sphere, but upon all others in space under necessary modifications, and as being carried on, from first to last, here and elsewhere, under immediate favour of the creative will or energy." A later expression of this view is that the alleged inadequacy of natural selection is said to make it impossible for evolution to be a self-acting mechanism, and must be directed by an external creative activity either Divine (Beibetz 1922-105, Barnes 1933:519-520) or psychical (Sampson 1909:13,29).

i) Restricted or Spiritualised Evolution. It is claimed that the theory of evolution cannot apply to man and/or the origin of the first life. In these respects there must have been supernatural intervention (J. Guibert 1900:136-146).

Today there is considerable pluralism amongst creationists. Historians agree that creationists are unable to agree on a unified theory (Numbers 1982, 1987, Barker 1985,1987). Ramm (1954), Thurman (1977:120-126) and Bradley and Olsen (unpublished report) have reviewed some of the more recent creationist theories. These include:

a) Literalism or "Creation-science". Life is believed to have been created recently in six days of twenty-four hours (Gish, Bliss and Bird 1981).

b) Flood geology is held by, for example, Whitcomb and Morris (1961:273-274).

- c) The Gap Theory has been defended by Custance (1976). One variant of the Gap Theory is that of Local Creation.
- d) The Alternate Day view, another variant of the the Gap Theory, holds that the geological ages come between twenty- four hour days of special creation.
- e) One interpretation of the Alternate Day view is that of Restricted Evolution. Evolution cannot apply, it is said, to matter, life and man which were specially created (Berry 1975:37).
- f) The Age-Day view, a species of Progressive Creation, in which the the days of Genesis are said to be eras corresponding to the geological record (Ross 1979:11).
- g) The Relative Time view. According to this view the days of Genesis were "twenty-four in another time context, but are years in our time context" (Bradley and Olsen).
- h) Variants of Gosse's Apparent Age view are held in part, for example, by Morris (1976:65,66). He suggests that the stars were created along with rays stretching to the earth, but rejects (somewhat inconsistently I think), Gosse's views.
- i) Narrational Day view. The days are twenty-four hours of revelation and recording of the creation story (Wiseman 1956).
- j) Directed Evolution has been defended by Appleby (1961).

While not all of the views above are mutually exclusive, it is evident that there has been no appreciable reduction in the diversity of thought amongst creationists from the turn of the century until now. With more than a century to come to an agreement, creationists still have widely diverging views over the time period for creation. It varies from around ten thousand years to billions of years. The creationists also differ widely over the extent to which they allow

supernatural intervention in natural processes. Some (e.g. Anderson and Coffin 1977) hold that all the major phyla were specially created and accept only micro-evolution within these phyla. For others (e.g. Berry 1975:37) special creation is reserved only for the appearance of matter, life and man. Still others (e.g. Appleby 1961) allow no special creation at all. This pattern of history suggests that creationism is pseudo-scientific.

Even within one viewpoint there can be widely differing explanations given for particular phenomena. While all "creation- scientists" agree that the earth is about 10 000 years old, they give very different explanations in the extremely difficult task (for them) of trying to account for the sequence in the fossil record. One would expect from the Literal View, that the fossils would not be laid down in any particular order and certainly not from the more primitive to the more complex.

Whitcomb and Morris (1961:273-274) give three sorting mechanisms purportedly explaining the fossil sequence. All of these involve flood geology. One is hydrodynamic sorting. The idea here is that the denser and more streamlined objects would have descended more rapidly and should be found in the lowest rock strata. Another is termed ecological sorting. The idea here is that the inhabitants of low regions would be overcome first by the flood waters and will be in the lowest strata. Another is anatomical sorting. The idea here is that the least mobile animals would have succumbed to the ravages of the flood first, and would be in the lowest strata.

Now the interesting thing is that all these positions have been continuously, and simultaneously held in spite of the fact that the fossil record contradicts them, as Gould (1983) has pointed out, and in spite of the fact that they are mutually incompatible. Whales, which are streamlined and marine, are found only in the upper strata, contradicting both the first and the second

explanations. Sloths, which are not known for their mobility are also found only in the highest strata contradicting the third explanation.

Another interesting case of diversity in thought amongst creation-scientists concerns their attempt to deal with the great distances of the stars from this earth. Most stars are more than 10 000 light years away. If the earth is only 10 000 years old, our nights should not be very starlit, but they are. One response is that the speed of light has slowed down since the time of creation, although not all creationists are agreed on this (Aardsma 1988). Morris (1976:65,66) proposes that the stars were created along with rays of light stretching to the earth.

Creationism in the twentieth century shows considerable diversity of thought. Even within particular viewpoints there is great diversity. Its lack of exclusionary convergence suggests that it is a pseudo-science. This is not to say that the pattern of history of creationism is the only indicator of demarcation, or that it is the root cause of its pseudo-scientific nature. Many other criteria, with some success, have been used to evaluate creationism (Kitcher 1982, Root-Bernstein 1983). Rather, its pattern of history is symptomatic of a deeper problem. Creationist theories are constrained more by their own traditions than by the real world. This has in many cases been quite candidly admitted.

Whitcomb and Morris (1961:xx) state that one of the purposes of their book is to... "examine the anthropological, geological, hydrological and other implications of the flood, seeking if possible to orient the data of these sciences within this biblical framework. If this means substantial modification of the principles of uniformity and evolution which currently control the interpretation of these data, then so be it."

Morris says (quoted in Root-Bernstein, 1983) that "no geological difficulties, *real* or imagined, can be allowed to take precedence over the clear statements and necessary inferences of Scripture" (emphasis mine).

Creationists would want to give the impression of being scientific. We should not expect this sort of admission to occur often. Occasionally, and very revealingly, their attitude to scientific data slips out. In one instance, Morris (1966:75-76) reviews an article by a fellow creationist and with disapproval says: "One gathers, however, that these beliefs could easily be forfeited if the scientific evidence should become strong enough to warrant it."

It is this type of attitude which has led to the pseudo- scientific character of creationism. I submit that the basic fault of creationism, of the type that seeks to be scientific, is that its theories are constrained by their interpretation of the Bible, and are only partially constrained by the real world. As a consequence its history shows no exclusionary convergence, and this pattern correctly reveals that it is a pseudo-science.

CHAPTER 3: MULTI-CORNERED TESTING: THE METHOD OF THEORY COMPARISON¹

Introduction

Disputes in Population Biology often centre on method. In such cases the controversy is not so much over what the facts are, but how they should be interpreted. What Tauber and Tauber (1977a) take as evidence for sympatric speciation, Hendrickson (1978) finds inconclusive. What Mayr (1963:17) takes as evidence for the Isolation Concept, Paterson (1981) finds unsatisfactory. Paterson (1982) has reported the need "for a more consistent logic if we are to reduce dispute and disagreement and improve understanding and comprehension in evolutionary biology".

How should we go about science? Philosophers of science do not have a unanimous answer (Oldroyd 1986:370). For some this is not even the right question to ask. Feyerabend (1975) says that there is no such thing as a set of inviolable methodological rules. Chalmers (1982:132) agrees. Others (e.g. Newton-Smith 1981, Gower 1986) dispute their methodological relativism. I will not enter the debate at this level. I will assume that some methods of doing science are better than others.

There is also dispute over which is better method: whether deductivism is superior to inductivism, or even whether these are different aspects of the same basic method (Popper 1972a:27-33, Newton-Smith 1981, Harré 1972, Andrewartha and Birch 1984). Neither will I enter the debate at this level.

Instead, my main purpose is to show that within a deductivist perspective there are three distinct procedures. These are verificationism, in which a hypothesis is supported by seeking to confirm it; falsificationism, in which a hypothesis is tested by trying to refute it; and multi-cornered testing (MCT), in which rival hypotheses are forced to confront each other².

Popper has supported falsificationism on the grounds that its underlying logic is valid. This is in contrast to verificationism. I will argue that these procedures can be evaluated apart from the validity or otherwise of their underlying arguments, and according to how much they tell us. From the perspective of this criterion, it will be argued that MCT is preferable.

The concept of MCT is well known to philosophers of science and scientists. The early Greek philosophers emphasized the struggle between alternate views (Feyerabend 1970:211). Mill (1910:111,112-114) saw the advantage of the "collision of adverse opinions." Chamberlin (1897) urged the use of what he called the method of multiple working hypotheses rather than the method of seeking to confirm a ruling theory. He was anxious that the scientist did not allow one theory to so dominate his mind that he fail to consider alternatives. Kuhn (1962:147) has said that it "makes a great deal of sense to ask which of two actual and competing theories fits the facts better." Popper (1972a:32-35, 1976:149) has emphasized competition between rivals as a means of testing. In sophisticated methodological falsificationism, an elaboration of the original philosophy by his followers, this is made more explicit (Weimer 1979). Lakatos (1970:115) has said that "...tests are - at least - three-cornered fights between rival theories and experiment...". Pannenburg (1976:56) and Laudan (1981b:145) agree that scientific method is, or should be, comparative.

Success in biology is frequently attributed to the use of this method. For instance, Platt (1964) says that the success of molecular biology is due to hypotheses being allowed to compete with one another. According to Gould

(1976) the honey-bee dance- language controversy was only resolved after the use of this method. Darwin's success was due, in part, to his comparison of the relative merits of Descent with Modification and Independent Creation (Ghiselin 1969:62-63). Success in the physical sciences has also been attributed to this method (Treitel 1987).

However, multi-cornered testing has been coupled with discredited notions about method in the past, and has been conflated with other methods. My one purpose is to show that the concept can stand without these discredited notions, and to emphasize the distinctiveness and superiority of MCT over, at least, the naive versions of verificationism and falsificationism. One reason for doing this is to try and bridge the communication chasm that exists between the philosophy of science and the practice of science (Gale 1984).

Towards a criterion for evaluating method

It has long been known that verificationist methodology hinges upon an invalid argument called the fallacy of affirming the consequent. For example:

If X is a dog it will have 4 digits on each forelimb.

X does have 4 digits on each forelimb.

Therefore X is a dog.

Compare this invalid argument with the following *modus tollens* argument:

If X is a dog it will have 4 digits on each forelimb.

X does not have 4 digits on each forelimb.

Therefore X is not a dog

Popper based his falsificationism on the grounds that the latter argument is valid. The problem with this account is that it fails to provide a role for confirmation whereas as Lakatos (1970:115) pointed out: "... some of the most interesting experiments result, *prima facie* in confirmation rather than falsification." In response, Popper (1976) has distanced himself from naive falsificationism arguing that growth in scientific knowledge occurs best through competition between rivals. I reach the same conclusion by looking at the problem from a different angle; by asking whether the confirmationist and falsificationist arguments can be evaluated apart from the validity/invalidity criterion. I think they can.

The first step is to convert the confirmationist argument into a valid one by altering the conclusion thus:

If X is a dog it will have 4 digits on each forelimb.

X does have 4 digits on each forelimb.

Therefore X could be a dog.

Since both the modified confirmationist argument and the *modus tollens* version are valid we can ignore the validity/invalidity distinction and look for other differences. What immediately comes to mind is that the modified argument tells us relatively little. It tells us that the dog hypothesis is a

possibility out of a finite number of possibilities, but this is not saying much. The *modus tollens* argument tells us more. While it fails to tell us what X is, it does tell us that X is not a dog. (We shall see later, however, that at the level of method we cannot conclusively exclude any alternative, because of the possibility of bringing in an *ad hoc* hypothesis). The important point is that an argument need not only be judged by the validity/invalidity criterion, but instead by how much it tells us.

The falsificationist argument, however, is not the best available. Only one of perhaps many alternatives has been excluded. A preferable method would be one which excluded many. An example illustrating the logic of this method, which I will call multi-cornered testing is as follows:

If X is a dog it will have 4 digits on each forelimb.

If X is a cat it will have 5 digits on each forelimb.

If X is a sloth it will have 2 digits on each forelimb.

X is either a dog or a cat or a sloth.

X has 2 digits on each forelimb.

Therefore X is a sloth

The logic of MCT is valid since it includes a premise to exclude all possibilities but three. (We shall see later that this premise is unrealistic). More interesting than its validity is that the LM argument tells us more than the previous two arguments. For on the basis of a single observation (line 5) we have excluded 2 alternatives and confirmed the other.

The feature that distinguishes MCT from the other two arguments is that it is comparative. For naive verificationism and naive falsificationism the issue is whether a particular theory squares with the evidence or not. For MCT the issue

is which theory fits the evidence best. This feature is not always made explicit. Shaw (1981:37-39) describes the comparative approach under the heading of affirming the consequent. The argument is justified because it is advanced against a background of unstated competing theories. My contention is that greater clarity will result in science if what is left unstated is made explicit. I have sought to show that multi-cornered testing deserves to be considered in its own right; it is distinct and better than the other two methods because the underlying argument tells us more.

The implication of the comparative approach is that there are distinct kinds of evidence, something which was appreciated by Francis Bacon if not earlier (Grunbaum 1976, Losee 1980:65), and recognised today (e.g. Glymour 1983:14). The one type will be called common evidence and will be contrasted with discriminatory evidence.

Common evidence is that kind which is equally favourable to two or more hypotheses. For example if we need to identify whether some animal X is a dog or a cat, it does not help to determine whether X has fur or not. 'Presence of fur' is common to both the dog and the cat hypotheses. 'Four digits on each forelimb', on the other hand is discriminatory evidence because it would tend³ to favour the dog hypothesis over its rival. Some important principles are involved in this distinction:

a) What is needed to support a hypothesis adequately is not just the *amount* of evidence but a certain *kind* of evidence. No amount of common evidence will support one hypothesis over its rivals.

b). Common and discriminatory evidence are relative terms. Whether 'presence of fur' constitutes common evidence depends on the context. In the case of deciding whether an animal is a dog or a cat, 'presence of fur' is common

evidence. If, however, the question is whether an animal is a mammal or a reptile, then 'presence of fur' is discriminatory evidence.

Unfortunately, as I have intimated, the situation is not so simple. The logic underlying MCT is clear and valid. However, as I will show in the next section, if one were to attempt to apply the argument to scientific practice a number of difficulties would arise. We shall see that this will force us to modify the logical structure of MCT given above (p.45). The aim is a methodological framework more useful for the practicing scientist.

The problem of additional unforeseen explanations

This difficulty is well known (Ghiselin 1974:4). X in the argument above could be a camel since camels as well as sloths have 2 toes on each forelimb. It is then unrealistic to limit the alternatives to 3, as statement 4 does in the MCT argument above. To solve this problem one could express the argument as follows:

If A_1 then B_1

If A_2 then B_2

⋮
⋮
⋮

If A_N then B_N

B_J

Therefore A_J - verification as $N - T$

(As N approaches T , hypothesis A_J approaches verification.)

where T = total number of possible hypotheses.

and $T \neq \infty$

N = actual number of hypotheses considered.

AJ, BJ = particular hypotheses and their predictions⁴.

Note that the argument remains realistic because it does not demand that $N=T$, something no one would want to claim certainty about. In so doing the argument becomes less than valid. It does not claim that AJ is the correct explanation, only that the closer one comes to considering all explanations the better confirmed AJ will be. Thus it gives a role to confirmation without demanding conclusiveness.

The crucial experiment and MCT share the same basic form in that both involve contests between rival theories and experiment. MCT differs in that it does not pretend that a single test will bring conclusive results. The best we can do as Hull (1967) puts it, is to aim for "successive approximation" to the truth.

The problem of ad hocery

A second difficulty, admitted by Popper, is that strict falsification is not possible because a "refuted" theory can be saved using an *ad hoc* hypothesis (Popper 1972a:42). He therefore distinguished between logic and method. At the level of method he provided the ruling that we should not try systematically to save our theories (Popper 1972a:78-86). The same problem faces MCT. Even if in a particular case one could prove that $N=$, one could still not conclude that one hypothesis has been vindicated over its rivals. In the biological example above one could save the dog hypothesis by suggesting that animal X has had two of its digits on each fore-limb amputated. The problem arises, as is well known, because it is not single statements that are tested in science, but "bundles" of statements (Duhem 1968:171-177, Quine 1953:41, Pannenburg 1976:56, Kitcher 1982:44, Ariew 1984). In our biological example we are testing at least two statements together- ' X is a dog' and ' X has the usual number of toes for the species'.

Whereas Popper says we should not try systematically to save our theories, it could be merely stipulated that the hypothesis be subjected to the same kind of test:

If X is a 2-toed dog it will have surgical scars.

If X is a sloth it will not.

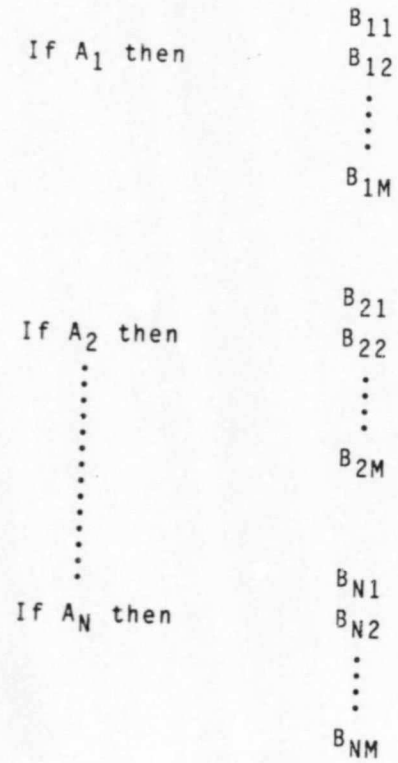
X does not.

Therefore X is a sloth.

Even now the conclusion is not conclusive but at least we have made demands on the dog hypothesis. It could be saved by the suggestion that the dog was born with an abnormal number of toes.

Despite this, the dog hypothesis is beginning to look rather poor- like phlogiston⁵ theory. It seems that for one hypothesis to begin to establish itself over its rivals, it must be subjected to a series of independent multi-cornered tests. The importance of independent evidence is well known (Ghiselin 1969:19-20, Paterson 1982:119). Now X might have been a dog. It is perhaps only from hindsight that the *ad hoc* nature of this hypothesis can be known. It is dangerous to attempt to say, *a priori*, how many tests will be required, for a theory may be dismissed prematurely. For the same reason, it is dangerous to give the ruling that we should not try to save our theories. It is not needed. As long as we continually force theories to compete, the demise of poor rivals will inevitably follow.

To indicate the necessity of a series of independent tests we can modify the original MCT argument thus:

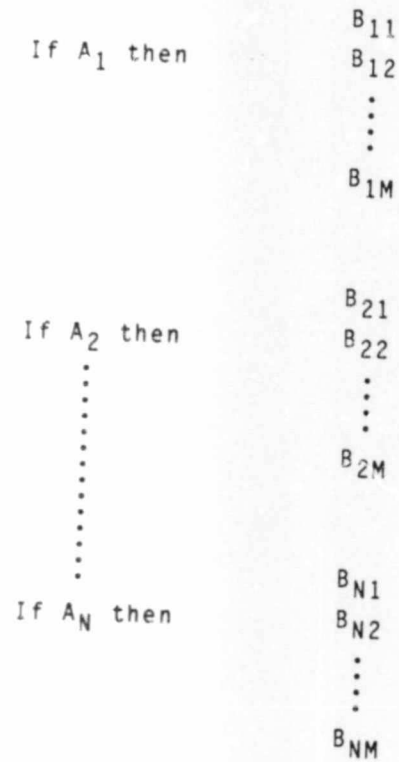


$B_{j1}, B_{j2}, B_{j3} \dots B_{jM}$

Therefore $A_j \rightarrow$ verification as $N \rightarrow T$ with $\uparrow M$

where $B_{j1}, B_{j2} \dots B_{jM} = M$ number of predictions of particular hypothesis A_j (other symbols as before).

Figure 3-1. The structure of multi-cornered testing.



$B_{j1}, B_{j2}, B_{j3} \dots B_{jM}$

Therefore $A_j \rightarrow$ verification as $N \rightarrow T$ with $\uparrow M$

where $B_{j1}, B_{j2} \dots B_{jM} = M$ number of predictions of particular hypothesis A_j (other symbols as before).

Figure 3-1. The structure of multi-cornered testing.

The problem of universality

There is a fourth difficulty in applying MCT to scientific practice. The hypotheses in the arguments above are spatio-temporally singular, i.e. of the form 'X is a dog.' However, scientific hypotheses are generally universal statements, i.e. of the form 'all dogs have 4 digits on each forelimb.' Now a single instance can confirm a spatio-temporally singular statement, but this is not so for universal statements. Finding that a particular dog has 4 digits on each forelimb does not confirm the hypothesis that all dogs have 4 digits on each forelimb. A comparative approach to method helps us to deal with this problem. To illustrate this consider the following arguments:

If all dogs have 4 digits on each forelimb then so will X.

X has 4 digits on each forelimb.

Therefore all dogs have 4 digits on each forelimb.

This argument is clearly invalid. It is always possible that some three-toed dog has so far escaped our attention. Our concern is not with validity but with learning as much as possible from an argument. To this end we give the above argument a comparative structure:

If all dogs have 3 digits on each forelimb then so will X.

If all dogs have 4 digits on each forelimb then so will X.

If all dogs have n digits on each forelimb then so will X.

X has 4 digits on each forelimb.

Therefore all dogs have 4 digits on each forelimb.

This argument is again invalid. However, we have excluded two hypotheses and shown that the one is the better of the three. To convert the argument into the valid form we can restructure it thus:

If A_1 then B_1

If A_2 then B_2

If A_3 then B_3

B_j

therefore A_j approaches verification

Now the argument has a similar form to the one previously, but this obscures an important point. There are two senses in which a hypothesis might be confirmed and these must be vigorously contrasted. The one is in the inductive sense, i.e. each supportive instance increases the confirmation of the universal statement. This is problematic. Popper (1972b:387-395) has argued that since the domain of a universal statement is infinite no amount of evidence can increase the probability of the statement being true. The second sense in which a hypothesis is confirmed is free from this difficulty. It only takes one supportive instance to confirm the one hypothesis *over* its rivals. The conclusion in the

above argument uses confirmation in this latter sense. This point has an important implication.

It suggests that we should distinguish between negative and discriminatory evidence. Negative evidence is absence of evidence and might be thought to count against a hypothesis. For instance, a person might object to the A2 hypothesis above on the grounds that Mongolian dogs have not been examined and these might have more or less than 4 digits on their forelimbs. But negative evidence could just as easily be brought against all rival hypotheses. Whatever the value of negative evidence, it has no value for deciding between rival hypotheses. For this purpose we need discriminatory evidence.

The problem of complementarity

It sometimes happens that two hypotheses, initially thought to conflict, come to be accepted as complementary rather than contradictory⁶. This should not, I think, deter us from applying MCT. Evidence that two apparently conflicting hypotheses are not mutually exclusive would be provided by the failure of the one hypothesis to establish itself over another, after a series of multi-cornered tests. It is not easy to tell *a priori* whether two hypotheses are actually conflicting. They may integrate at a point higher than our present level of understanding.

The problem of the theory dependence of observation

The logical positivists thought that observation preceded theory. Several philosophers of science have argued that observation always requires a theoretical framework (Kuhn 1970b:242, Popper 1970:51). Popper (1972a:47) says that these theoretical frameworks are themselves guided by observations, but these observations are in turn guided by a frame of reference. The

relationship between observation and theory seems to be complex (Chalmers 1982:22-34, Newton-Smith 1981:19-43, Weckert 1986) involving a self-correction process (Hesse 1980:108). A good way of describing this relationship is to say that observation and theory are interdependent (Sparkes 1981). A case for interdependency can be made by referring back to the biological example of an MCT argument:

If X is a dog it will have 4 digits on each forelimb

If X is a cat it will have 5 digits on each forelimb

If X is a sloth it will have 2 digits on each forelimb

X has 2 digits on each forelimb

therefore it is a sloth.

The first three statements can be divided into two parts: the hypothesis and the prediction. The fourth statement is an observation statement. It is quite easy to demonstrate that theory precedes observation from this example. The observation (statement 4) only comes after the hypotheses and their predictions have been made. This is so because we need the hypotheses and predictions to tell us what to observe- 'number of digits' rather than 'colour of X' (say).

This biological argument can also be used to demonstrate the sense in which observation precedes theory. Our hypotheses are not wild guesses. They have emerged from some unstated observations (and theories). Note that the possibility of X being a snake or an earthworm is not even considered. Why? Because we have already observed that X has fur and limbs which snakes and earthworms do not have. The test was begun with the exclusion of certain hypotheses; this was done on the basis of previous observation and theory.

We have seen the sense in which the observation statement (line 4) is dependent on theory. But this can be exaggerated. The observation statement is also subject to the dictates of the real world. (I am assuming the existence of such a thing). One cannot count more digits than there are. The interrelationships between theory, observation and the real world are depicted in figure 3-2

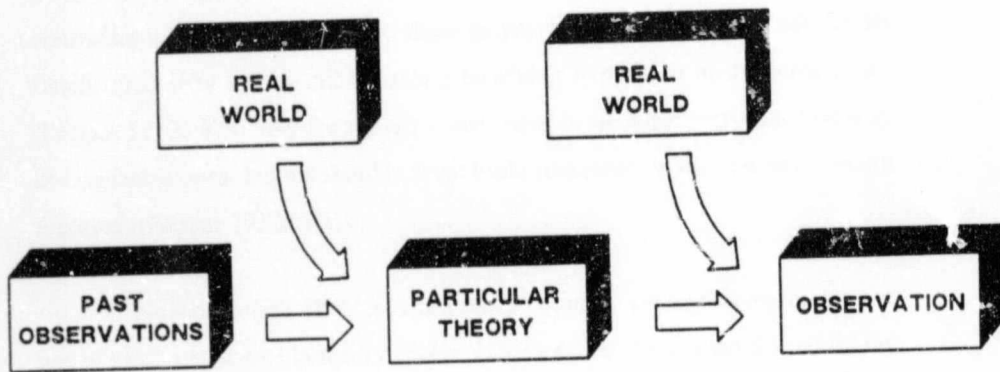


Figure 3-2. The relationship between observation, theory and the real world.

The problem of the status of basic statements

An example of an observation statement or basic statement is 'X has 2 digits on each forelimb'. It is on the grounds of a basic statement that one chooses between alternative hypotheses. Its status then is very important. Popper (1972a:105) claims that basic statements can never be justified. They have the character of dogmas. Acceptance of them is a matter of convention. In order to legitimize the use of basic statements in theory choice, Popper has provided certain methodological rules. Examples are: "If accepted basic statements contradict a theory, then we take them as providing sufficient grounds for its falsification only if they corroborate a falsifying hypothesis at the same time" (Popper 1972a:87). "We should not accept stray basic statements- i.e. logically disconnected ones- but we should accept basic statements in the course of testing theories" (Popper 1972a:106).

Newton-Smith (1981:63) considers Popper's attempt to legitimize the use of basic statements in theory choice to have failed. He argued that under the Popperian scheme, acceptance of basic statements are arbitrary decisions. No matter how many basic statements one accepts, the collective decision is still arbitrary.

For Deutscher (1968) Popper's non-justificationist stance on basic statements means that scientific theories could never actually be falsified. He says that for a theory to be falsified we must know or have good reason to believe that a basic statement is true; what we need is justified or at least partially justified basic statements.

The earlier claim that there is a sense in which observation is theory dependent seems to support a partial justificationist account of the status of basic statements. I cannot develop the theme here, but let me suggest the possibility that the status of basic statements lies somewhere between being dogmas and being justified; that they are indeed not justified because they are theory

dependent, but neither are they dogmas because they are subject to the dictates of the real world.

Have we supplied basic statements with enough cognitive weight to legitimize their use in theory choice? Science has been described as a negative feedback system (Sparkes 1981:157). It is one of the properties of such a system that it can tolerate imperfections because it is self-correcting. If this is so, then it suggests that science can work with less than justified basic statements, as long as they are better than dogmas.

Conclusions

My aim has been to show that three basic procedures can be distinguished within a deductivist perspective, this is not apparent if the arguments underlying these procedures are evaluated only according to the validity/invalidity distinction. This was necessary because these procedures have been conflated in practice. MCT is distinct because it is comparative. Of the three, it is preferable because it tells us the most.

It seems advisable to couple the following with this approach: the recognition that a multi-cornered test does not produce absolutely conclusive results; a reliance on a series of independent tests; an emphasis on the interdependency between observation and theory and the recognition that basic statements are higher than dogmas.

In the next section, I divide MCT into several steps as a guide to its application in science.

Applying MCT

The following steps are suggested:

1. Recognize common evidence.
2. Define the alternative hypotheses distinctly.
3. Develop multi-cornered predictions from the alternatives.
4. Decide between the alternatives on the basis of discriminatory evidence.

Step 1 involves admitting that whereas in the past a particular body of data had been used to support a single hypotheses, it is in fact consistent with two or more hypotheses. This is, in fact, always a possibility. Philosophers of science agree that "more than one theoretical construction can always be placed on a given collection of data" (Kuhn 1962:76). In step 2 the hypotheses are distinctly defined such that $H_1 \neq H_2 \neq \dots H_n$. This is an important step in that it makes testing between the alternatives possible (Pannenburg 1976:67). Paterson (1981) has drawn attention to the fact that the failure to appreciate the difference between different species concepts is a major problem making the testing of hypotheses very difficult. In step 3 predictions such that $P_1 \neq P_2 \neq \dots P_n$ are made from these hypotheses. In step 4 the discriminatory evidence corresponding to a particular prediction P_j is used to adjudicate between the original hypotheses.

It must be noted that this sequence is somewhat artificial since the steps are dependent upon one another and not necessarily in the order given. For example, in clarifying an hypothesis it may become evident that it actually consists of two mutually exclusive hypotheses. Loehle (1987) says that a period of theory maturation is needed to determine the consequences of a theory. Note too, that the steps given above are for a single multi- cornered test. In an actual research programme a number of these tests may be run concurrently.

Notes

1 The core of this chapter was published in Anderson (1987).

2 More sophisticated versions of verificationism and falsificationism involve competition between theories. The contrast here is between MCT and the naive versions.

3 I say *tends to* because of the possibility of ad hocery. Some evidence initially thought to be discriminatory can be made consistent with a rival through the addition of an *ad hoc* hypothesis.

4 or (loosely) deductions. I do not necessarily mean predictions in the forward-looking sense.

5 According to this theory, a substance called phlogiston escapes from metal whenever it is combusted. Later Lavoisier showed that metal weighs more after combustion and not less as was expected from phlogiston theory. Some attempted to save the theory by suggesting that it had negative weight (Hempel 1966:29-30).

6 The classic case is quantum field theory, in which the wave and particulate natures of light are resolved as complementary hypotheses in a single theory (Whitehead 1926:213-214, Gardner 1981).

CHAPTER 4: METHODOLOGY AND THE SPECIES PROBLEM

Introduction

Species biology is known for its polemics and pluralism to the point that Paterson (1981) has called for the examination of the foundations of species biology. He wrote that "...in science, inadequate concepts and inconsistent logic often lie at the root of disputation and disagreement." This chapter is written in the belief that as Mayr (1976a:500) put it "the species problem is a topic where productive collaboration between philosophy and empirical biology is possible." The main object will be to evaluate the actual practices of biologists according to the criteria established in chapter two. In addition I hope to show that the problem of pluralism in species biology can be avoided through applying the multi-cornered method; and to suggest further methodological principles of benefit to species biology.

The field is a large one. I cannot hope to discuss all the definitions of species given in the literature. This project will be restricted to the three major concepts found in the literature under the assumption that if the normative methodological principles apply to these they will apply universally. First, I will describe these concepts.

The three major species concepts

Mayr (1976a) has distinguished between the Morphological and the Biological Species Concepts. According to the Morphological Species Concept

"a species is a group of individuals or populations with the same or similar morphological characters (Mayr 1942:115, 1976c)." The Biological Species Concept in contrast is genetical, species are defined in terms of their biological characteristics such as reproductive isolation and interbreeding. Paterson (1981) describes two quite distinct genetical concepts:

a) The Isolation concept: Species are groups of natural populations reproductively isolated from other groups. Paterson calls this concept a relational one because the species is defined with reference to other species.

b) The Recognition concept: A species is "that most inclusive population of individual biparental organisms which share a common fertilization system" (Paterson 1985). This concept is genetical but not relational. Originally Paterson's concept read: "members of a species share a common mate recognition system." The mate recognition system is a subset of the fertilization system. As a consequence the later definition is more general.

I turn now to the evaluation of method in species biology. The method used by biologists will be compared to the steps involved in the multi-cornered approach.

Step 1: Recognizing common evidence

Mayr (1963:31), although he did not use this term, noted that proponents of the morphospecies concept relied on common evidence. He describes how the concept is typically justified: "...natural populations considered by general consent to be species are morphologically distinct. Morphological distinctness is thus the decisive criterion of species rank. Consequently, any natural population that is morphologically distinct must be recognized as a separate species (Mayr 1963:31)."

Dobzhansky (1951:261-2) agrees that this is the way the concept is supported: "there has been no shortage of attempts to define a species. The lack of morphological intergrades between species has frequently been depended upon..."

In simplified form the argument is:

If the morphological species concept is primary then species would be
found to be morphologically distinct.

Species are in fact found to be morphologically distinct.

Therefore the morphological species concept is primary.

Mayr brings out that the conclusion is fallacious because it overlooks the strictly secondary role of morphological differences. The primary criterion of species rank of a natural population according to him is reproductive isolation. Our morphologists argument is inadequate because it fails to consider the possibility of an alternative and more primary concept equally accounting for the morphological differences between species. In other words, morphological distinctness is not recognized to be common evidence.

It is interesting that, while Mayr rejects the morphospecies concept, he uses the same fallacious argument to support the Isolation Concept. The underlying theme is the obvious match between this concept and the observation that species are reproductively isolated. In the following quotes we will find that reproductive isolation, which, from the perspective of the multi- cornered approach is common evidence, is taken as supporting the Isolation or Biological Species Concept.

Mayr (1963:316) writes that "... the term species signifies a very definite mutual relation between sympatric populations namely that of reproductive isolation."

And:

"...This striking discontinuity between sympatric populations is the basis for the [biological] species concept in biology" (Mayr 1963:17).

And again:

"Indeed the most characteristic attribute of a species in such a non-dimensional system is that it is separated by a gap from other units in the system. The gap that surrounds a species is the core of the species concept" (Mayr 1963:18).

Although it is variously expressed, the underlying argument is the same:

If the Isolation Concept is primary then species would be found to be
reproductively isolated.

Species are in fact reproductively isolated from other species.

Therefore the Isolation Concept is primary.

What is interesting about this argument is that it is not multi-cornered. The possibility that reproductive isolation is common evidence, and that there is an alternative hypothesis (e.g. the Recognition Concept), that could equally account for this evidence is not considered. In the next section I assess whether species biologists have practiced the second step in multi-cornered testing.

Of course, both genetical concepts of species predict that there will be a lack of intergrades between species, so that this is common evidence.

Step 2: Distinctly defining the alternatives

This is an important step in the multi-cornered approach. Before one can adequately test between alternatives it is necessary to distinguish clearly between them. Attention has been drawn to the fact that the failure to appreciate the difference between different species concepts is a major problem making the testing of hypotheses very difficult (Paterson and Macnamara 1984). Elsewhere Paterson (1981) writes: "Concepts of species are found in two distinct fields of biological endeavour, Taxonomy (the science of classification), and Population Genetics. In both these fields distinct concepts are used and their conflation is a major source of confusion". Mayr (1976b:485) says that "after a thorough study of most of the literature on this subject I have come to the conclusion that the disagreement is due to the fact that there are in existence not merely one but actually three entirely different species concepts. All past arguments and discussions have suffered from the fact that an author has often championed one of the concepts against the others or has wavered between two of these without realizing it."

Judging by the amount of conflation occurring in the literature it is necessary to ask whether step 2 has been practiced by species biologists. The question is whether the various species definitions in the literature have been adequately delimited. First, it will be necessary to discuss some normative principles on the subject of defining. To begin with a simple case, my object will be to answer the question 'how should we define gold?' Now gold is a spatiotemporally unrestricted class, whereas it has been argued that species are individuals, i.e. spatiotemporally localized entities (Hull 1976, Ghiselin 1974). Despite this difference I will argue that many of the principles derived from the simpler case can be carried over to the species problem, even if they have to be

appropriately qualified. In particular, it will be argued that species definitions should be based on that characteristic considered to be the most important biologically.

Normative principles in defining

According to the traditional theory of meaning it is the conjunction of properties associated with a term that determines the set of all those entities to which the term correctly applies (Kripke 1977). For 'gold' the properties are yellow, heavy, fusible, etc. Anything which has these properties *is* gold. The intension of the term is said to determine its extension. This theory is problematic.

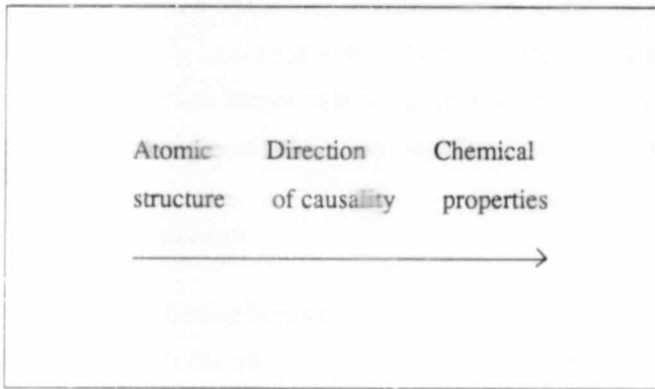
One can imagine the possibility that a substance has the properties normally associated with gold and yet not be in fact gold. The substance could be fools gold. The traditional theory of meaning would demand that we regard it as gold. There are several possible responses to this problem. Kripke (1977) suggests one¹. Another would be to enlarge our list of properties so as to exclude fools gold. The problem here is that our definition of gold is likely to become unwieldy. A third possibility which will be explored here is to distinguish between superficial and underlying properties of gold, and to use the latter in defining. Within the traditional theory of meaning then, it is possible to distinguish two ways of defining.

Two ways of defining

a) Gold could be defined in terms of its superficial properties. e.g. its colour, fusibility etc. or

b) Gold could be defined in terms of its underlying properties. e.g. its atomic structure (indicated by its atomic number).

Gold's atomic structure and superficial properties can be placed along a causal chain. The atomic structure of gold determines its chemical properties. Gold is yellow because of its particular atomic structure. We do not want to say it the other way round: that gold has its atomic structure because it is yellow. There is, then, a particular direction of causality. This can be represented thus:



The methods of defining given immediately above are based on different events in the causal chain. To define in terms of superficial properties is to define in terms of effects. To define in terms of underlying properties is to define in terms of causes. The two methods will henceforth be called 'defining by causes' and 'defining by effects'. In the next section, these two methods will be evaluated. The criterion that will be used is universality. It is generally agreed that universality is an important requirement for a law to be properly called scientific. Let us consider each method in turn.

Defining by effects

To determine how universal a definition is, it is necessary to express it in terms of a conditional. A conditional is a statement which allows one to speak of potential as opposed to actual events. The conditional would be: 'if 'x' has E_1 , E_2 , E_3 ,..... under condition C it is B' where E_1 - E_n are superficial properties.

To test for universality we substitute potential cases and determine if the statement would hold true under all possible conditions. Under certain circumstances, such as absolute zero temperature and the absence of light, a substance might have the atomic structure and yet would not have the external properties normally associated with it. Also, a substance might have superficial properties of gold and yet not be in fact gold. The substance could be fools gold. Defining by effects does not satisfy the requirement of universality.

Defining by causes

The conditional here would be: 'if 'x' has U_1 , U_2 , U_3 ,.... under condition C it is B' where U_1 - U_n are underlying or causal properties.

Definitions of this form do satisfy the requirement of universality. There is no conceivable instance where x would have atomic number 79 and yet not be gold. It makes no difference what conditions are specified. The conditional would still hold true under conditions of absolute zero temperature and no light.

In general, there is a connection between the degree to which a definition has been formulated using causes rather than effects and its universality. Since universality is a requirement of scientific laws, definitions derived from causes are preferable.

What about definitions that include both cause and effect? For example, 'anything which is yellow, heavy and has atomic number 79 is gold.' The

definition then becomes long and complicated, and implies that all these properties are equally significant. It seems reasonable to stipulate that a definition be in terms of that property which is deemed to be most significant.

Another reason why definitions by causes are preferable is that they are better candidates for dealing with the ontological rather than the epistemological question. We do not want to say that what makes an item a member of the class of gold items is its colour, fusibility etc., even though these are useful indicators of an item's true identity. It is more plausible to say that what makes an item a member of the class of gold items is its atomic structure.

From classes to individuals

A number of principles have emerged out of the above discussion. Definitions should be based on causes rather than effects and should be ontological and universal. It is not clear that these can be applied to the species problem. The difficulty is the nature of species. Some have said that species are sets or classes (Gregg 1950, Burma 1976, Kitcher 1984) and not individuals (Bernier 1984, Ruse 1987). Many believe that species are individuals (Ghiselin 1974, Sober 1984) and because they are individuals they do not have essences (Hull 1978, Ghiselin 1969, 1984) and thus "cannot be defined at all" (Hull 1976). Others (e.g. Kitts 1983) have agreed that species are individuals, but claim that despite this, they do have essences (in the Aristotelian not Platonic sense). In this discussion I am going to follow Kitts and assume that species do have an essence; and that they should be defined in terms of that essence.

It must be stressed that doing so is not to fall into typological thinking. Whether reproductive isolation or a common fertilization system is the essence of species, these are properties of populations not individuals.

Furthermore, even if species do not have essences, this does not mean we have to throw out all of the criteria above. Instead of demanding that definitions be ontological it could be merely stipulated that they be based on those characters which are most directly symptomatic of species. Hull (who is not an essentialist) agrees that "species must possess theoretically significant properties" (Hull 1984:111). It seems reasonable to stipulate that a definition must be in terms of these properties, even if we do not take it to be a definition in the traditional sense.

He says further that "... there is a close connection between the theoretical significance and support of counterfactual claims." Most would agree, then, that definitions if not ontological should be at least the most directly symptomatic of species, and should be universal. This, we shall see, is enough to adjudicate between the various species concepts.

Species concepts and definitions: an evaluation

My purpose will be to evaluate various species definitions according to the criteria established in the previous section. The need for this is evident from Eldredge's (1985) comment on the species problem in this context. He felt that "most discussions of the "species problem" have really been addressed to the the epistemological problems of species recognition and have led rather automatically to the conclusion that there are different "sorts" of species. Most such discussions never really address the core ontological question: Just what if anything *is* a species?" My concern will not be with the truth or falsity of the definitions as such. This question will be left for later. Indeed I want to argue that unless species definitions conform to certain rules it is difficult if not impossible to test them. One definition of the Biological Species Concept is "species are groups of interbreeding natural populations that are reproductively isolated from other such groups (Mayr 1969)." The important point about this concept is that it treats two events in the speciation process, interbreeding and reproductive

isolation, as if they were at the same causal level- as if they were of the same importance. He has not separated out which of the two features he considers to be definitive, or at least the most directly symptomatic, of species. This will not do. If the criteria established in the previous section are normative, Mayr must decide which event he takes to be prior and base his definition on that. If he believes isolation to be primary then he must give his definition as: species are natural populations reproductively isolated from other such groups.

As the definition stands it is ambiguous how the 'are' should be read. He could mean that reproductive isolation and interbreeding are the definitive characteristics of the species in the same way that gold's atomic structure is the definitive characteristic of this metal. Or he could mean that reproductive isolation and interbreeding are properties of the species as yellow and fusibility are to gold. If one is only concerned with practicality, there is no need to distinguish between underlying and superficial properties, so one would include a little bit of everything in the definition. Mayr sometimes seems to take this line. He writes (1976c:484): "a successful definition should not lay too much stress on any of the above listed criteria [morphological characters, genetic distinctness and lack of hybridization]." Of the biological Species Concept he says: "such a definition is applicable in practically all difficult cases of bird taxonomy" (Mayr 1976c:484). Yet he still maintains, somewhat inconsistently I think, that "reproductive isolation is the most important quality of the species" and that "reproductive isolation always remains the primary criterion" of species rank (Mayr 1957:381).

The trouble with Mayr's definition, then, is that it is too broad. Paterson puts it this way: "The term Biological Species Concept is not very descriptive - it's too wide. All genetical concepts of species are equally biological" (Paterson and Macnamara 1984). He went on to say that when the concept is used it conflates two sometimes three different concepts of species.

Including both cause and effect in the definition makes it empirically untestable. For, no-one denies that interbreeding and reproductive isolation are characteristics of species. If this is all that is meant, then the definition cannot be pitted against any other definition.

Other definitions of species suffer similar problems. For example, Littlejohn (1981:319) gives the following overly broad definition which he admits is lengthy. It begins: "A biparental species consists of a set of similarly adapted individuals, the members of which can combine genetic information through sexual reproduction." A description follows of the mechanisms which are supposed to allow the individuals to combine genetic information.

Mayr (1976c) has also given an expanded definition of species. From the perspective of the normative principles discussed above the shorter definition is better. The expanded definition reads: "A species consists of a group of populations which replace each other other geographically or ecologically and of which the the neighbouring ones intergrade or hybridize wherever they are in contact or which are potentially capable of doing so (with one or more of the populations) in those cases where contact is prevented by geographical or ecological barriers." This definition is less adequate because it includes three events which are at different causal levels - replacement, hybridization and isolation. The result is a broad and unnecessarily complicated definition.

I submit that definitions should not be merely practical; simply concerned with the problems of species identification. Instead they should attempt to address the question 'what is a species'. Definitions should make the claim that some single characteristic is the primary or definitive one.

Consider the Isolation concept. One way the Isolation concept might be interpreted is: "species are reproductively isolated from other species." As it stands this version of the Isolation concept is compatible with the Recognition

Author Anderson Michael Laurence

Name of thesis Aspects Of Scientific Methodology With Special Reference To Evolutionary Biology. 1988

PUBLISHER:

University of the Witwatersrand, Johannesburg

©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.