

species concerned is now extinct. Understandably, our taxonomist's managers have been unwilling to allow her to revisit her studies of genera A, B, X and Y -- at least until she has completed her current revision of the distantly related Genus Z which comprises some 300 or so species, many of which have been found to contain powerful antiviral agents potentially of immense economic importance.

Articles such as Ian's are so very difficult to write. On the one hand, there is a need for a simplified account of methodology that can be acted upon. On the other, one must avoid venturing too far into a specialised area without the necessary specialist equipment, thus running the risk of giving hostages to fortune either by losing the attention of one's readers or (worse still) by exposing one's own inevitably limited knowledge of the terrain, which has certainly changed a bit since Popper's time.

In conclusion, might it not be better to draw the attention of CALM's scientists to that excellent little book by A.F. Chalmers, namely *What is this thing called Science? An assessment of the nature and status of science and its methods* published by University of Queensland Press (1988) and currently available at a cost of \$17.95? Chalmers is readable, comprehensive and balanced. He covers everything from induction, through theory-dependent observation, falsification (Popper), theories as structures (Lakatos' Research Programmes, Khun's Paradigms), objectivism, anarchism (Feyerabend), realism, instrumentalism and truth, unrepresentative realism, and many other "isms".

Perhaps the more interesting issues raised by the taxonomic parable above are managerial rather than scientific in nature.

### Comment on "Use of hypotheses in Science and Information Division" by Ian Abbott

*Matthew Williams*

I enjoyed this article, and my only reservation about it is that it was too short. Because of this, the article oversimplified some aspects of the scientific method relating to hypothesis testing. In my comments below I try to clarify some of these aspects:

1. The article argues that the hypothetico-deductive method is the only method of science. This is wrong. Sir Ronald Fisher also stated (Fisher (1935): "The Design of Experiments") that the purpose of experimentation is to test hypotheses, and he was wrong too. Consider, for example, an experiment designed to compare five different tree species in their production of some commodity, say the amount of millable timber each produces after ten years. In a typical experiment, the five species would be planted together in a number of replicates, perhaps at a number of sites. The yield of timber would then be compared, and analysed, probably by ANOVA. The implicit null hypothesis underlying such an ANOVA is "(X): All five species produced the same amount of millable timber (i.e. the mean yields are all equal)" (this could be tested by the F statistic). However, nobody would ever believe such a null, and the purpose of this null is merely to provide a framework for the experiment. We do not even need to look at the F statistic, other than to assess the probability that any observed differences were due to chance. The real purpose of this experiment was "(Y): to objectively determine the best producing species, and to estimate its yield". It seems to me that to state hypothesis (X) as part of the aim of this experiment would be unnecessary (and indeed misleading), whereas (Y) should be stated as part of the aim.

2. Another point that I think needs clarification is the distinction between hypothesis testing and decision making. These two tasks, the former scientific and the latter political, are quite separate. In the discussion of the two alternate hypotheses (A: "Prescribed burning of jarrah forest in spring does not spread *Phytophthora* fungus"), and (B: "Prescribed burning of jarrah forest in spring causes the spread of *Phytophthora* fungus"), it is stated that "If the evidence collected does not refute (A), then we do not need to consider (B)". I agree that it is nonsensical to attempt to test (B), since (A) is a null hypothesis and hence more amenable to disproof. However, I think there are two points which should be noted:

(i) To assume that (A) is true, until disproved, is to confuse the artifice of the hypothetico-deductive technique with common sense. Before we test (A), we do not know if *Phytophthora* fungus is spread by spring burning. We merely propose (A) (as a null hypothesis) because it is a (relatively) easy statement to test using the hypothetico-deductive technique. [As an aside, how we choose to manage the forest until (A) is proved or disproved is a purely political decision - it may be prudent, in order to protect the forest resource, to assume that (B) is true and act accordingly.] If the evidence collected does not refute (A), then we should still consider (B). We must always bear in mind that if our experiment is inadequate (for example, too small), then we may never be able to refute (A). In the final decision of whether (A) or (B) is true, we need to consider the weight of evidence relating to (A) and (B). If a very small experiment fails to reject (A), our chance of making a type II error (i.e. accepting (A) as true when in fact it is false) is high. Conversely, this chance is low if a very large/comprehensive/powerful experiment was performed. [This fact has led to certain environmental protection agencies in the US requiring that in environmental impact studies, the power of the experiment designed to detect an impact, is specified. Otherwise, the outcome could be "rigged" to provide a result of no impact, simply by doing a very small study.]

In conclusion, in making decisions about the jarrah forest, managers need to consider both (A) and (B), and the chance that each is true. Thus it is perfectly defensible for managers to assume that (B) is true, if they adjudge that the evidence

as true when in fact it is false) is high. Conversely, this chance is low if a very large/comprehensive/powerful experiment was performed. [This fact has led to certain environmental protection agencies in the US requiring that in environmental impact studies, the power of the experiment designed to detect an impact, is specified. Otherwise, the outcome could be “rigged” to provide a result of no impact, simply by doing a very small study.]

In conclusion, in making decisions about the jarrah forest, managers need to consider both (A) and (B), and the chance that each is true. Thus it is perfectly defensible for managers to assume that (B) is true, if they adjudge that the evidence refuting (B) is weak. Just how much evidence is needed before (B) is refuted depends to some extent on how much damage the spread of *Phytophthora* fungus does, versus the gains or savings to be made from spring burning. In other words, a political decision is still needed, based upon the scientific evidence.

(ii) At the conclusion of the article, it is stated that the formulation of the null hypothesis “... seems to be analagous to the presumption in criminal law of innocence until guilt is proven.” I think it is worth pointing out that it is not necessarily reasonable to assume that because the presumption of innocence is used in criminal law, that it is a valuable assumption in all decision making situations. In law, the “presumption of innocence” (and associated “burden of proof” on the prosecution) is employed because society puts the cost of falsely imprisoning the innocent far above the cost of releasing the guilty. This is a value judgement, and if we as a society decide that it is far better to ensure the imprisonment of the guilty at the expense of punishing some innocents, it would be reasonable to reverse both the presumption of innocence and the burden of proof.

Thus, in the context of the hypotheses (A) and (B), if we consider spreading *Phytophthora* fungus very costly, it would be reasonable to assume that (B) is true until evidence disproving this is found. However, if we consider not spring burning more costly, we should perhaps assume that (A) is true until disproved.

In summary, I think it is necessary to keep clear the distinction between statistical and logical hypothesis testing, and decision making. Making decisions is all about value judgements and is thus a political activity. The hypethetico-deductive technique is simply one tool used to decide between alternate hypotheses. However, it is no substitute for “common sense” decision making.