Statistics as Rhetoric in Psychology

I.D. John
University of Adelaide

The confusion and misunderstanding of inferential statistics found amongst a class of exiting third-year psychology students is shown to parallel similar reported misunderstandings amongst the wider psychological community. The history of inferential statistics and of their institutionalisation within psychology is briefly described, and some suggested reasons for these developments are discussed. It is argued, supported by illustrative examples, that the rhetorical association of statistical inference with scientific method serves to assert and maintain epistemic authority in psychology.

Statistics have a particular or even unique importance in the psychology curriculum. In Australia where, in keeping with a tradition of neo-colonial cultural deference, the institutionalisation of psychology usually self-consciously closely follows United States models, it has been observed of the undergraduate psychology curriculum that "one component will certainly cover research methods, including psychological statistics" (Lovibond, 1977, p. 106), and that statistics is usually the only course "which everyone agrees should be taught and should be mandatory for all students." (McNicol, 1988, p. 279).

In the light of the consensus about the importance of statistics in psychological education, the level of understanding of students exiting from first degree courses might be thought to be a matter of some interest. Most such students will not undertake any further formal study of statistics, but will, it is hoped, be thoughtful and informed users of psychological research who are able critically to relate research findings to psychological knowledge claims. Those who do proceed to further study will be assumed to have established a sound grasp of the basic concepts required for more advanced work.

Some information gathered from one exiting third-year class hardly bears out these expectations. Asked to complete a questionnaire, they showed a consistent tendency to concur with propositions exaggerating the conclusions properly to be inferred from experimental evidence, and those able to calculate experimental effect sizes from appropriate information consistently overestimated their magnitude — under some conditions by a factor of more than 2.

There are no obvious reasons to believe that this finding would not be repeated amongst other groups of psychology students after 3 years of study, but in discussing it with a number of psychologists little concern has been expressed. It has been suggested that such interest as it might have lies in illustrating the limitations of individual cognitive processes and students' reliance on inappropriate and biasing heuristics for estimating effect size.

It is not surprising that psychologists should seek explanations in terms of individual psychological processes, but the inadequacy of an exclusively individualistic psychological explanation for this state of affairs may be illustrated by asking the reader to contemplate the likely consternation in psychology circles if, in contrast to what has been found in this case, it were to be revealed that as a result of their education, psychology students were consistently disposed to underestimate the size of experimental effects and the evidentiary value of research findings.

Rather than explore this situation in terms of individual psychological processes, this paper will attempt to relate it to the broader context in which it occurs, and to take it as a starting point for a consideration of the way in which inferential statistics serve a rhetorical purpose in providing epistemic authority in psychology. From this perspective, what is most noteworthy about the students' distorted and exaggerated ideas about the evidentiary value of statistics is not so much that they exist, but that they engender so little concern and seem to be tacitly condoned. They are, indeed, largely unexceptional within the context of the statistical discourse and practices which have become institutionalised within psychology.

MISUNDERSTANDING OF STATISTICS

There is considerable evidence of widespread misunderstanding and exaggeration of the evidentiary value of statistics within psychology. Oakes (1986) looked at the beliefs held by 70 British academic psychologists about the logic of significance tests, and concluded that 67 of them "had a less than sound grasp of the conclusions that may properly be drawn from the results of a significance test" (p. 82). In additional studies he examined psychologists' intuitions about measures of effect size, such as the proportion misclassified and the correlation coefficient, and found a "tendency heavily to overestimate the degree of relationship that has actually been demonstrated" (p. 91). A similar tendency, amongst psychology graduate students and undergraduates, to overestimate effect size associated with particular values of t and specified Ns (sample size) has been reported by Posavac and Sinacore (1984).

Oakes further argued "that significance tests encourage the widespread but manifestly untenable belief that hypothetical experimental findings assume the mantle of reality at the arbitrary 0.05 level of significance" (p. 85), thus corroborating Rosenthal and Gaito's (1963) claim of a cliff effect, in the interpretation of levels of statistical significance, which takes the form of a relatively precipitous loss of confidence in moving from the .05 to the .10 level.

Confusion between psychological significance and statistical significance has also been noted in an analysis of 50 reports sampled from the British psychological literature (Cochran & Duffy, 1974). This study found a relatively high incidence of incorrect use of statistics (24%) and misinterpretation of the results of hypothesis testing (22%). Cochran and Duffy were, however, primarily concerned with the adequacy of sampling procedures and concluded from an examination of 276 studies that 85% of them probably used inadequate sampling, and of these only 5% discussed the possible implications of this for the results of the research. Grichting (1989) conducted a similar study based on 1273 articles, published over a 10-year period in major Australian psychology and sociology journals, and reached comparable conclusions. In his view, significance testing was badly abused and a disproportionately large number of studies were based on inadequate sampling procedures which precluded statistical inferences to clearly defined populations. Rosnow and Rosenthal (1989) concluded after examining 320 published articles that in less than 1% of cases employing analysis of variance were interaction effects correctly interpreted.

Requests for reprints should be sent to I.D. John, Department of Psychology, University of Adelaide, GPO BOX 498, Adelaide SA 5001, Australia.

In the light of this level of misunderstanding and confusion it is not surprising that disagreement about the statistical interpretation of individual studies should also occur. Peters and Ceci (1982) observed, in their well-known study of previously published articles resubmitted for publication, that "[p]erhaps the most serious objections that reviewers had about the previously accepted manuscripts were directed towards the studies' designs and statistical analyses" (p. 190). Weitzman (1984) claimed that two experiments used as exemplary illustrations, in a methodological analysis published in the *Psychological Bulletin* (Cook, Gruder, Hennigan, & Flay, 1979), demonstrated seven separate pitfalls in the statistical analysis which invalidated any conclusions drawn from the data, and Schwartz and Dalgleish (1982) have argued that the data from an experiment by Mueller and Wherry (1982) warrant a conclusion entirely opposite to that originally advanced.

It should not be thought that this confusion about the proper use of statistical inference is restricted to psychology, since similar concerns have been raised about the use of statistics in other social sciences (e.g., Gould, 1970; Skipper, Guenther, & Nass, 1967) and in biomedical research (Bailar, 1986; Gore, Jones, & Rytter, 1977), but what should make this situation a matter of particular concern is the focal role that statistics have assumed in the authorisation of psychological knowledge claims.

INSTITUTIONALISATION OF STATISTICS

Danziger (1985) has observed that:

> "psychology appears to be unique in the degree to which statistical inference has come to dominate the investigation of *theoretically* postulated relationships. In this discipline it is generally assumed without question that the only valid way to test theoretical claims is by the use of statistical inference ... The methodology has become highly institutionalized, providing important criteria for publication policies and scientific reputations. Faith in this methodology certainly unites a much larger number of research psychologists than does any kind of commitment to a particular theoretical framework. It is surely the most serious candidate for the status of a generally accepted puzzle solving paradigm in modern psychology" (p. 3)

In being naturalised within psychology, inferential statistics have become taken for granted as a universal, coherent, noncontroversial collection of rule-governed algorithms for the mechanisation of the production of conclusive knowledge, despite a history of continuous unresolved controversies over contradictory and irreconcilable philosophical and theoretical positions (Danziger, 1987; Gigerenzer, 1987; Gigerenzer & Murray, 1987; Gigerenzer et al., 1989).

History

Fisherian statistics, which had been developed for use in agricultural research after World War I, were first used in psychology in the late 1930s (Rucci & Tweney, 1980). Practical questions addressed by research workers and administrators in education were not dissimilar to those encountered by agricultural scientists — does a particular manipulation, say the application of a fertiliser or the introduction of a new teaching technique, have a reliable effect on crop yield, in the first case, or educational attainment in the second? This meant that the potential application of analysis of variance and complex experimental designs to practical problems in education was readily grasped and from there their use spread to psychology.

Fisher, however, interpreted his procedure of null hypothesis testing, not as a means for practical decision-making, but as a solution to the problem of induction and a rigorous method of scientific inference. In this belief, inferential statistics were rapidly and enthusiastically taken up by psychologists as a means for the testing and choice of theories. Within psychology, the previously independent Wundtian experimental tradition of investigation of the psychological processes of the individual had been merged with the Galtonian tradition of description of the characteristics of large populations in a new neo-Galtonian model (Danziger, 1987). This was based on the assumption, firmly rejected by those psychologists operating from a clinical or dynamic perspective outside of the experimental tradition, that, in examining psychological processes, the distribution of responses of a group of subjects could be treated as equivalent to the distribution of repeated responses of the individual subject, thus allowing statistical inferences to underlying individual theoretical processes from group data.

After the Second World War the theoretical position of Neyman and Pearson became more widely known. In contrast to Fisher, with whom they were engaged in heated controversy, they saw inferential statistics as a procedure for making a decision about which of two specified hypotheses was best supported by a set of observations. They introduced the idea of statistical power, which has no meaning within the Fisherian system, as a fundamental concept and required the prior specification of probability levels. The concept of statistical power underlines the difficulty of drawing conclusions from a failure to reject a null hypothesis, and the potential importance of effect size as opposed to statistical significance level in the interpretation of data. Notwithstanding the irreconcilable differences between the two opposing theoretical positions these were largely ignored within psychology, where an anonymous hybrid theory emphasising the importance of statistical significance above all else became institutionalised.

> "Well known textbook authors and editors of journals put forward numerous erroneous beliefs, such as that level of significance by itself determines (1) the magnitude of the effect, (2) the probability that the null hypothesis is true or false, (3) the probability that the alternative hypothesis is true or false, and (4) the degree of confidence that the experimental result is repeatable" (Gigerenzer et al., 1989, p. 209)

THE CRITIQUE

The institutionalisation of inferential statistics in psychology has taken place notwithstanding opposition and criticism represented by such views as:

> "I believe that the almost universal reliance on merely refuting the null hypothesis as the standard method for corroborating substantive theories in the soft areas is a terrible mistake, is basically unsound, poor scientific strategy, and one of the worst things that ever happened in the history of psychology" (Meehl, 1978, p. 187).
>
> "Statistical significance testing uses a corrupt form of the scientific method. Even if properly used in the scientific method, educational research would still be better off without statistical significance testing" (Carver, 1978, pp. 397-398).
>
> "The arguments ... surely present an irrefutable case for the abandonment of routine significance testing" (Oakes, 1986, p. 66).

Concerns which, from the outset, were voiced about the misuse of tests of statistical significance, such as those discussed in Morrison and Henkel's (1970) *The Significance Test Controversy*, have been largely ignored and, it has been alleged, have even encountered unusual difficulty in being published (Bradley, 1984). When the shortcomings of a particular statistical practice, such as disregarding the limited power of many tests, has been unequivocally demonstrated, the warning has gone largely unheeded (Sedlmeier & Gigerenzer, 1989; Rossi, 1990). A substantial number of practitioners, who might be expected to be eager consumers of psychological knowledge, find the information provided by the application of tests of statistical significance irrelevant to their needs (Barlow, Hayes & Nelson, 1984; Morrow-Bradley & Elliot, 1986) and, it has been claimed, the statistics curriculum in US graduate schools fails to equip students to tackle many types of research problems of current interest (Aiken, West, Sechrest, & Reno, 1990).

A basic consideration which has been lost sight of, or suppressed, in the institutionalisation of inferential statistics in psychology is that a point null hypothesis is almost invariably false, and given a sufficiently large sample and sufficiently small error variance can almost invariably be rejected. Conversely, given an insufficiently large sample and a large enough error variance it can not be rejected. It follows that the results of an inferential test are, in themselves, ambiguous and their interpretation is dependent on usually unexamined underlying assumptions. This under-

lies the importance of personal judgment and discretion in interpreting tests and illustrates why inferential statistics can not properly be applied mechanically, or regarded as rule-governed algorithms for the production of psychological knowledge.

SUGGESTED EXPLANATIONS

If, as seems likely, such criticism in other areas of psychology would have undermined and rendered unsupportable any possible claims for paradigmatic or foundational status, it remains to be explained how inferential statistics continue to occupy such a central position in an enterprise which publicly represents itself as committed to the discovery, by rational means, of conclusive and indubitable psychological truths. Numbers of answers have been suggested to this question.

Weitzman (1984) has suggested that since a null hypothesis is almost invariably false there is a high probability of being able to reject it even when an intervention is ineffective, and so null hypothesis testing attains the strength of an abundantly reinforced habit. Further reinforcement is likely to follow as a consequence of "a bias amongst editors and reviewers for publishing *almost exclusively* studies that reject the null hypothesis via statistical significance testing" (Kupfersmid, 1988, p. 637).

Probability values seem to provide a convenient alternative metric for experimental effects, which are originally often expressed in units of measurement that are difficult to interpret in terms of practical importance (Carver, 1978).

Education and training play a role in perpetuating the present situation. Numbers of arguments raised against inferential statistics in the historical debate closely parallel the type of objections raised by beginning students who frequently feel, as can be verified by any teacher, that what they are being taught is counterintuitive (Lunt & Livingstone, 1989; Oakes, 1986). The teaching of statistics largely ignores basic philosophical issues pertaining to knowledge production (Dar, 1987) and, in distinction to other areas of psychology, avoids study of the central theoretical controversies and disputes in favour of instruction in the cookbook application of various statistical techniques, so that students come unthinkingly to apply tests of statistical inference routinely as a kind of knowledge increase ritual.

However important these factors might be it does not seem plausible that they can, either individually or together, explain the focal position which inferential statistics have assumed in the process of psychological knowledge production. If we are to believe, as is maintained by the standard view, that this is an objective and rational process by means of which the true nature of reality is reliably revealed, the situation outlined seems at least anomalous or contradictory and perhaps even hypocritical.

In the opinion of some observers it appears to constitute a self-serving form of collective, primary process thinking. Widely held beliefs have been spoken of as "fantasies" (Carver, 1978) and "illusions" (Dar, 1987), "which psychologists seem not to wish to be cured of" (Gigerenzer & Murray, 1987, p. 25), and which permit the presentation of ideology as scientific fact (Atkins & Jarrett, 1979). They create an air of scientific (Dar, 1987) and philosophical respectability (Oakes, 1986), "and are necessary to maintain the dream of mechanized inductive inference" (Gigerenzer & Murray, 1987, p. 25)

AN ALTERNATIVE ANALYSIS

Epistemic Authority

As a way out of the dilemma posed by this situation, an alternative to the view of science usually accepted in psychology may be considered. Pierre Bourdieu, the French sociologist of science, claims that:

> "The scientific field is the locus of a competitive struggle in which the *specific* issue at stake is the monopoly of *scientific authority*, defined inseparably as technical capacity and social power, or, to put it another way, the monopoly of *scientific competence*, in the sense of a particular agent's socially recognised capacity to speak and act legitimately (i.e. in an authorised and authoritative way) in scientific matters." (1975, p. 19)

That is, he sees science as characterised by a struggle for epistemic authority, which, once attained, may be applied by the scientist to any number of personal and social ends, rather than by an objective and rational method which guarantees the indubitability of the conclusions it reaches about the nature of reality.

In highly autonomous scientific fields, such as are found in the natural sciences, epistemic authority is conferred relatively independently of the wider community by a socially recognised and legitimated group of competing knowledge producers who, as a consequence of their own contributions to knowledge, are deemed to be cognitively competent to grant recognition to others. In the social sciences the situation is more complicated:

> "... because the power which is at stake in the internal struggle for scientific authority within the field of the social sciences, i.e. the power to produce, impose and inculcate the legitimate representation of the social world, is one of the things at stake in the struggle between the classes in the political field. It follows that positions in the internal struggle can never attain the degree of independence in relation to positions in the external struggle which is to be found in the natural sciences." (Bourdieu, 1975, p. 36)

Psychologists are, in the main, keenly sensitive and even resentful about the way in which they must compete for epistemic authority, in psychological matters, with lay outsiders who, in their view, are technically incompetent. Eysenck (1957) dismisses their views as "meretricious sesquipedalianism" and Skinner (1976) has been equally scornful:

> "The disastrous results of common sense in the management of human behavior are evident in every walk of life, from international affairs to the care of a baby, and we shall continue to be inept in all these fields until a scientific analysis clarifies the advantages of a more effective technology. It will then be obvious that the results are due to more than common sense." (Skinner, 1976, p. 258)

These attitudes help to explain psychologists' highly developed self-consciousness about their scientificity and their "passionate interest in the natural sciences; what is at stake via their claim to impose the legitimate definition of the most legitimate form of science, i.e. natural science, in the name of epistemology or the sociology of science, is the definition of the principles of evaluation of their own practice." (Bourdieu, 1975, p. 24).

Scientific Knowledge

Associated with the idea of science involving a struggle for epistemic authority is a position about the nature of scientific knowledge. The test of scientific knowledge from this perspective is not its correspondence with the nature of reality, since such a test is, in principle, incapable of being applied, but the extent to which it commands agreement amongst the relevant knowledge community. The justification of a scientific proposition is not accomplished by demonstrating that it correctly describes reality, but in demonstrating, either by doing or talking, that there are convincing reasons warranting agreement with it. A person is conceded to know something if they can reliably cause an event to happen at will, or bring it under their control, or if they can advance reasons which convince the authorised members of the relevant knowledge community. In the physical and natural sciences, occurrences such as space voyaging and splicing new genes into the genome of an organism are spectacular demonstrations of knowing, by comparison with which the events which psychologists can bring under control, such as the serial position effect in verbal rote learning, are likely to seem very modest indeed. In psychology and the social sciences, in the relative absence of compelling demonstrations of knowing, the production of knowledge is much more reliant on purely discursive means such as negotiation, argument, and persuasion.

Summary

To summarise: The psychological enterprise, like other social institutions and groups, aspires to a position of privilege in society and does so on the basis of its claim to produce, apply, and command an authoritative body of scientific knowledge. The interests of the enterprise are integrally bound up with the assertion and acceptance in the wider community of its epistemic authority. This task is difficult, firstly, because authorised psy-

chological knowledge claims about the nature of reality are likely to be in competition with beliefs which serve to justify the interest of other important social groups, and, secondly, because psychologists have had little success in demonstrating their knowledge by bringing significant events under their control in an impressive way. It is to be expected, therefore, that psychological knowledge claims will be more keenly contested and disputed than those of the natural sciences and that the form and conduct of argument in psychology will be correspondingly more critical.

THE REVIVAL OF INTEREST IN RHETORIC

A growing appreciation of the essentially contested nature of scientific knowledge has seen a surge of publications in recent years discussing the place of rhetoric in science in general (e.g., Campbell, 1975; Overington, 1977; Schuster & Yeo, 1986; Weimer, 1977), in the human sciences, (e.g., Gusfield, 1976; McCloskey, 1983; Nelson, Megill, & McCloskey, 1987; Perelman, 1979; Simons, 1989; Weigert, 1970), and in psychology in particular (Bazerman, 1987; Billig, 1987, 1990).

To acknowledge the importance of rhetoric in science and in psychology is not to endorse a distinction which opposes rhetoric to a purportedly objective, logical, and rational scientific method; that itself is a frequently encountered rhetorical tactic. It is to acknowledge "that scientific argument is essentially persuasive argument and therefore is rightly termed *rhetorical* in the sense defined by students of 'the new rhetoric', where 'rhetoric' denotes the entire field of discursive structures and strategies used to render arguments persuasive in given situations" (Schuster & Yeo, 1986, p. xii).

Rhetoric is a nonformal type of reasoning for arriving at, or justifying, conclusions. It is to be distinguished from formal inference which, commencing with explicit premisses, permits the drawing of certain or apodictic conclusions, and, exemplified by syllogistic reasoning, is usually taken as a model of rationality. Most reasoning, of necessity, falls short of arriving at certain or final conclusions and involves not justifying the truth of conclusions, but justifying adherence to them. The test of the conclusion is not its correspondence with some state of the world, but the sense of conviction which it engenders in reasoners.

Argumentative persuasion, or rhetoric, involves a wide and complex range of techniques, but in essence "The argumentative process consists in establishing a link by which acceptance, or adherence, is passed from one element to another." (Perelman, 1979, p. 18). The process is ubiquitous and is to be seen in advertising, where the esteem associated with culturally valued icons is passed to commercial products, in Melanesian cargo cults where an envied ideal of material wealth is linked via a bamboo simulacrum of a radio antenna to the cult leader's prophecies of "cargo", and in psychology where prestigious, but only partially assimilated, philosophical and metatheoretical ideas about the indubitable and conclusive nature of scientific knowledge are transferred in psychological discourse to particular institutionalised methods and practices of knowledge production.

RHETORIC IN PSYCHOLOGY

Psychological Texts

This can be illustrated by an examination of psychological texts and, particularly, research reports, which play a key role in the production and dissemination of psychological knowledge claims and are the most institutionally regulated form of psychological discourse. Bazerman (1987) has argued that the style of scientific reporting which is laid down in detail in the *Publication Manual of the American Psychological Association* embodies rhetorical conventions consistent with behaviourist assumptions. "The behaviorist method then could be considered identical to scientific method, excluding other forms of psychological investigation as unscientific. And the behaviorist rhetoric could be identified as the only proper way to write science" (pp. 134–135).

The nature of the literary persona, or the individual who is speaking to us through the impersonal conventions of the psychological research report, is revealing in this regard. His or her identity and institutional affiliation are announced at the commencement of the report, but thereafter the author disappears from view and his or her role is limited to holding up a figurative window (Gusfield, 1976) through which the reader is invited to look at reality, and ultimately, by following the same sequence of rational steps, to confirm independently the exact conclusions which have been reached by the author. The author, by his or her absence, is projected as an objective, detached, and undeviating follower of methodological prescriptions — an automaton who preferably might be replaced by a computer, and whose personal characteristics are not so much irrelevant as dangerous if allowed to intrude.

A central feature of the research report is the attempt to establish epistemic authority by a stress on adherence to methodological canons. Such an appeal to scientific method as an incontrovertible and final justification for knowledge is a powerful persuasive strategy in a society which regards science as the major source of epistemic authority. This type of account also resonates with the mechanistic root metaphor which underlies the standard or behaviourist position in psychology. If the human operates like a machine, psychological knowledge producers can hardly be treated as an exception. However, the contrived and conventionalised nature of the research report is disclosed by the obvious and undisguised passion with which authors, in everyday life, identify and defend their priority claims for the knowledge they have produced.

Statistics and the Appeal to Scientific Method

The appeal to scientific method is the major persuasive tactic used in psychological argument in the struggle for epistemic authority. The foundational status of scientific method is emphasised in psychological discourse, but the nature of the method is never made explicit. This is because the demarcation between science and nonscience is strongly disputed and none of the suggestions so far advanced about the distinctive and defining characteristics of the scientific method such as induction, hypothetico-deduction, or falsification have proved tenable. Nevertheless, psychological discourse strongly implies some inescapable connection between inferential statistics and scientific method and exploits the common belief in a status hierarchy of the sciences in which scientificity is equated with quantification and the use of mathematics.

This is convincingly illustrated by recent *Proposed Guidelines for the Preparation and Publicity of APS* [Australian Psychological Society] *Information and Publicity Material* (1986). Members and agents of the Society are informed that "Information and publicity material ... must conform to the following guidelines" (p. 25), one of which has to do with psychology as a science. It requires that "Material directed to intending students of psychology must stress the scientific nature of the discipline, ... its use of computers for data storage and reduction, and its reliance on statistical methods for the evaluation of data" (p. 25). Apart from its imperious tone, the proposed guideline is noteworthy for the way in which it equates statistical methods with the scientific nature of the discipline and stresses the importance of the mechanisation of knowledge production by computer. Another example of the rhetorical linking of epistemic authority and statistics comes from an editorial in a newsletter for psychological practitioners. "Psychologists are experts in cognitive functioning. We are trained in the use of statistics and understanding the notions of probability and reliability." (Editorial, 1986). Prominent publicists for psychology such as Eysenck employ similar arguments, "psychology, being a science, uses methods and discussions of a highly technical nature. In particular, it uses statistical methods." (Eysenck, 1957, p. 17).

The rhetorical nature of this type of argument about the use of statistics for the production of authoritative psychological knowledge is underlined by the selectivity of the use of this tactic. Many psychologists who accept statistical arguments as a justification for personally congenial conclusions are unwilling to accept these same arguments as a justification for conclusions about contentious issues to which they are opposed.

The overriding rhetorical strategy in psychology of valorising inferential statistics by an appeal to their scientificity, and by representing them as exemplifications of the scientific method, is supported by other subsidiary arguments based on partially assimilated philosophical and metatheoretical ideas. These are frequently encountered in both explicit and implicit forms; they do not constitute a coherent and consistent whole, but are drawn on as resources in particular argumentative situations. Contrary to what might be suggested by psychological theories stressing the importance of cognitive consistency and balance, consistency is less important than having access to cognitive resources which meet the immediate demands for justifying conclusions in particular argumentative situations. Some examples of these arguments may be considered.

Supporting Arguments

The standard perspective in psychology combines aspects of empiricism and realism, and envisages a world constructed of facts which it is the task of psychologists to discover. It is often argued as if statistical significance testified to facticity. This is a powerful tactic in the socialisation of students, whose experience of psychological data gathering often falls short of the controlled and systematic ideal, and may be somewhat messy and chaotic. Nevertheless, to be assured, at the end of the day, that the application of a statistical test has transubstantiated their often rather ragged data into a statistically significant finding, or a psychological fact is, for many students, an impressive and somewhat mystifying experience. In the absence of a detailed technical and theoretical background from which to draw counter arguments, agreement is coerced and doubters are likely to feel intimidated about openly expressing their reservations.

The belief that facts are discovered also underlies and is used to justify the practice, increasingly common with the freer access to computers, of indiscriminately using inferential statistics to sift through large masses of data. The statistically significant relationships which emerge are then accorded the status of scientifically discovered facts, which by some process of Baconian induction, will hopefully lead to the formulation of appropriate theory. The enthusiasm for inferential statistics leads to the paradoxical situation where the legitimate data reducing, descriptive, and summarising role of statistics is lost sight of, and modest inputs of data are transformed into greatly amplified volumes of computer printout, which are eagerly but uncritically scanned for any signs of statistically significant relationships. The vastly increased computational capacity which computers have made available to students and researchers has often not been accompanied by comparable increases in basic understanding.

Sophisticated inductivists early abandoned the possibility of arriving at certain knowledge inductively, but continued to entertain a watered down version of inductivism and the prospect of being able to specify the probability of scientific propositions. Superficial acquaintance with these views and with Bayesian theory seems to be mobilised in justifying the frequently encountered perception of probability values as inverse probabilities, referring to the truth of an hypothesis, rather than referring to the conditional probability of the data under the null hypothesis.

Psychologists, nowadays, usually have at least some passing familiarity with Sir Karl Popper's views about the crucial role of falsification in the process of scientific knowledge production, and often mistakenly feel that this is consistent with, and justifies the practice of, null hypothesis testing. If a null hypothesis is rejected, the next step in the inferential chain is to accept or treat as confirmed its converse, the experimental hypothesis. However, from a Popperian perspective confirmation is only a weak basis for theory choice. The possibility that an experimental hypothesis might ever be rejected by null hypothesis testing, and thus convincingly provide the type of evidence sought by Popperians, involves considerations of statistical power and is highly problematic.

CONCLUSION

In conclusion, two matters remain to be addressed. The first is the question of what can replace the role that significance testing currently has in the production of psychological knowledge. This question itself reflects the very epistemological stance and particular supposition that there are algorithms, or canonical procedures, for knowledge production which has been the target of criticism in this discussion. However longingly we may envisage the prospect of such procedures and wish to be unburdened of our uncertainty, they have, until now, proven to be a mirage, and the onus is surely on those who postulate their existence to make the case for them.

The second is the *tu quoque* response which this discussion will predictably provoke. The rhetorical nature of the arguments advanced is freely acknowledged. To do otherwise would be to contradict the basic thesis of this discussion, and to implicitly endorse the position which has been criticised, namely that there are logically flawless, conclusive, nonrhetorical means of psychological knowledge production of which statistical inference is the most important example.

It is unlikely that psychologists will be readily disposed to abandon that position since it is a powerful strategy in their struggle for epistemic authority. Abandonment of the idea that inferential statistics are rule-governed algorithms for the production of apodictic conclusions would undermine the epistemic status claimed for much psychological knowledge and reduce the distinction between "scientific" and other currently less privileged, or excluded, kinds of psychological knowledge. The illusory prospect of indubitability and conclusiveness is likely to continue to be more attractive than the cognitive discomfort occasioned by confronting the abiding uncertainty of our knowledge claims. However, to do this could assist the self-understanding of psychologists and help render intelligible aspects of the process of psychological knowledge production which might otherwise seem opaque.

REFERENCES

Aiken, L.S., West, S.G., Sechrest, L., & Reno, R.R. (1990). Graduate training in statistics, methodology and measurement in psychology. *American Psychologist, 45,* 721–734.

Atkins, L., & Jarrett, D. (1979). The significance of "significance tests". In J. Irvine, I. Miles, & J. Evans (Eds.), *Demystifying social statistics* (pp. 87–109). London: Pluto.

Bailar, J.C. (1986). Science, statistics and deception. *Annals of Internal Medicine, 104,* 259–260.

Barlow, D., Hays, S., & Nelson, R. (1984). *The scientist practitioner: Research and accountability in clinical and education settings.* New York: Pergamon.

Bazerman, C. (1987). Codifying the social scientific style: The APA *Publication Manual* as a behaviourist rhetoric. In J. Nelson, A. Megill & D.N. McCloskey (Eds.), *The rhetoric of the human sciences: Language and argument in scholarship and public affairs* (pp. 125–144). Madison, WI: University of Wisconsin Press.

Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology.* Cambridge: Cambridge U.P.

Billig, M. (1990). Rhetoric of social psychology. In I. Parker & J. Shotter (Eds.), *Deconstructing social psychology* (pp. 47–60). London: Routledge.

Bradley, J.V. (1984). Antinonrobustness: A case study in the sociology of science. *Bulletin of the Psychonomic Society, 22,* 463–466.

Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. *Social science information, 14,* 19–47.

Campbell, P.N. (1975). The *personae* of scientific discourse. *Quarterly Jounal of Speech, 61,* 391–405.

Carver, R.P. (1978). The case against statistical significancce testing. *Harvard Educational Review, 48,* 378–399.

Cochran , R., & Duffy, J. (1974). Psychology and scientific method. *Bulletin of the British Psychological Society, 27,* 117–121.

Cook, T.D., Gruder, C.L., Hennigan, K.M., & Flay, B.R. (1979). History of the sleeper effect: Some logical pitfalls in accepting the null hypothesis. *Psychological Bulletin, 86,* 662–679.

Danziger, K. (1985). The methodological imperative in psychology. *Philosophy of the Social Sciences, 15,* 1–13.

Danziger, K. (1987). Statistical method and the historical development of research practice in American psychology. In L. Krüger, G. Gigerenzer, & M.S. Morgan (Eds.), *The probablistic revolution* (pp. 35–47). Cambridge, MA: MIT Press.

Dar, R. (1987). Another look at Meehl, Lakatos and the scientific practices of

psychologists. *American Psychologist, 42,* 145–151.

Editorial. (1986, December). *Newsletter of Private Practising Psychologists, 2.*

Eysenck, H.J. (1957). *Sense and nonsense in psychology.* Harmondsworth: Penguin.

Gigerenzer, G. (1987). Probablistic thinking and the fight against subjectivity. In L. Krüger, G. Gigerenzer, & M.S. Morgan (Eds.), *The probablistic revolution* (pp. 11–33). Cambridge, MA: MIT Press.

Gigerenzer, G., & Murray, D.J. (1987). *Cognition as intuitive statistics.* Hillsdale, NJ: Erlbaum.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life.* Cambridge, MA: Cambridge U.P.

Gore, S.M., Jones, I.G., & Rytter, E.C. (1977). Misuse of statistical methods: Critical assessment of articles in BMJ from January to March 1976. *British Medical Journal, 1,* 85–87.

Gould, P. (1970). Is *statistix inferens* the geographical name for a wild goose chase? *Economic Geography, 46,* 439–448.

Grichting, W.L. (1989). Psychology and sociology in Australia: The published evidence. *Australian Psychologist, 24,* 115–126.

Gusfield, J. (1976). The literary rhetoric of science: Comedy and pathos in drinking driver research. *American Sociological Review, 41,* 16–34.

Kupfersmid, J. (1988). Improving what is published: A model in search of an editor. *American Psychologist, 43,* 635–642.

Lovibond, S.H. (1977). Undergraduate teaching in psychology in Australian Universities. In M. Nixon & R. Taft (Eds.), *Psychology in Australia: Achievements and prospects* (pp. 100–117). Sydney: Pergamon.

Lunt, P.K., & Livingstone, S.M. (1989). Psychology and statistics: Testing the opposite of the idea you first thought of. *The Psychologist, 12,* 528–531.

McCloskey, D.N. (1983). The rhetoric of economics. *Journal of Economic Literature, 21,* 481–517.

McNicol, D. (1988). Psychology and the higher education system. *Australian Psychologist, 23,* 273–288.

Meehl, P.E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.

Morrison, D., & Henkel, R.E. (Eds.). (1970). *The significance test controversy: A reader.* Chicago: Aldine.

Morrow-Bradley, C., & Elliot, R. (1986). Utilization of psychotherapy research by practicing psychologists. *American Psychologist, 41,* 188–197.

Mueller, J.H., & Wherry, K.L. (1982). Test anxiety and reaction time for matching decisions. *Journal of Research in Personality, 16,* 281–289.

Nelson, J.S., Megill, A., & McCloskey, D. (Eds.). (1987). *The rhetoric of the human sciences: Language and argument in scholarship and public affairs.* Madison, WI: University of Wisconsin Press.

Oakes, M. (1986). *Statistical inference: A commentary for the social and behavioural sciences.* Chichester: Wiley.

Overington, M.A. (1977). The scientific community as audience: Toward a rhetorical analysis of science. *Philosophy and Rhetoric, 10,* 143–165.

Perelman, C. (1979). *The new rhetoric and the humanities: Essays on rhetoric and its applications.* Dordrecht: Reidel.

Peters, D.P., & Ceci, S.J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *The Behavioral and Brain Sciences, 5,* 187–255.

Posavac, E.J., & Sinacore, J.M. (1984). Improving the understanding of statistical significance: Reporting effect size. *Knowledge: Creation, Diffusion, Utilization, 4,* 503–508.

Proposed guidelines for the preparation and publication of APS information and publicity material. (1986). *Bulletin of the Australian Psychological Society, 8,* 25.

Rosenthal, R., & Gaito, J. (1963). The interpretation of levels of significance by psychological researchers. *Journal of Psychology, 55,* 33–38.

Rosnow, R.L., & Rosenthal, R. (1989). Definition and interpretation of correlation effects. *Psychological Bulletin, 105,* 143–146.

Rossi, J.S. (1990). Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting and Clinical Psychology, 58,* 646–656.

Rucci, A.J., & Tweney, R.D. (1980). Analysis of variance and the "second discipline" of scientific psychology: A historical account. *Psychological Bulletin, 87,* 166–184.

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin, 105,* 309–316.

Schuster, J.A., & Yeo, R.R. (Eds.). (1986). *The politics and rhetoric of the scientific method.* Dordrecht: Reidel.

Schwartz, S., & Dalgleish, L. (1982). Statistical inference in personality research. *Journal of Research in Personality, 16,* 290–302.

Simons, H.W. (Ed.). (1989). *Rhetoric in the human science.* London: Sage.

Skinner, B.F. (1976). *About behaviorism.* New York: Vintage.

Skipper, J.K., Guenther, A.L., & Nass, G. (1967). The sacredness of .05: A note concerning the uses of statistical levels of significance in social science. *American Sociologist, 2,* 16–18.

Weigert, A.J. (1970). The immoral rhetoric of scientific sociology. *American Sociologist, 5,* 111–119.

Weimer, W.B. (1977). Science as a rhetorical transaction: Toward a nonjustificational conception of rhetoric. *Philosophy and rhetoric, 10,* 1–29.

Weitzman, R.A. (1984). Seven treacherous pitfalls of statistics illustrated. *Psychological Reports, 54,* 355–363.