Phil. Soc. Sci. 15 (1985) 1-13

The Methodological Imperative in Psychology*

KURT DANZIGER, *York University, Toronto*

1. METHOD AND THEORY

If we enjoy contemplating science as though it were a finished edifice we can limit ourselves to its theories and observations. But if we regard it rather as an ongoing work of construction we cannot really ignore the scaffolding of procedures.[1] Methodical procedure produces the observations that count as scientific, but at the same time it is the repository of explicit and implicit theoretical assumptions. The relationship between observation and theory is mediated in practice by methodological prescriptions.

'Valid' theory, 'acceptable' observations, and 'appropriate' methodology are enmeshed in relations of mutual interdependence. In addition to the theory-dependence of observations that arises from the nature of human communication there is a theory-dependence of observations which is mediated by methodological rules.[2] If theoretical preconceptions are an unavoidable component of methodological rules, and if such rules mean that only certain kinds of observation will ever be made, then a certain predetermination of observation by theory must follow. We have here the possibility of a *methodological circle* where methods based on assumptions about the nature of the subject matter only produce observations which must confirm these assumptions. Within such a circle theoretical change would be limited to the set of theories which share the assumptions incorporated in the methodological rules. Any theoretical change beyond this would have to involve a methodological change.

The point is that the methods used to test a theory may presuppose the truth (or falsity) of the theory to be tested. Investigators commonly play with loaded dice. Not all kinds of outcomes have an equal chance of appearing in their research situations. This renders the application of probability theory to experimental outcomes problematic, because the

* *Received* 12.8.83

1 L. Laudan, *Progress and its Problems: Towards a Theory of Scientific Growth*, Berkeley 1977, chap. 2.

2 G. Böhme, 'Die Bedeutung von Experimentalregeln für die Wissenschaft', *Zeitschrift für Soziologie*, 3, 1974, 5-17.

indifference principle is quite frequently violated by the methods employed. Of course, the violation will be a matter of degree, at times it will be insignificant and at other times extreme. The trouble is that we have no precise way of estimating this bias. Simply examining the outcomes of a series of studies will not do, because we cannot separate the contributions of the investigator's methods from the contributions of factors assumed to operate independently of the application of these methods. If we operate in a situation where a great deal of very precise knowledge already exists we may be able to arrive at a reasonable estimate of bias due to method, but in psychology this state of affairs typically exists only in very limited areas and in connection with theoretical generalizations of very specific application. For theories that claim a broader scope we lack the precision and for complex experimental situations we lack the knowledge that would be necessary to make a meaningful estimate of the degree of bias.

This issue becomes acute when we leave the level of the individual study and face the question of comparing the results of numerous studies that purportedly address the same topic. In psychology empirical topic areas typically consist of a collection of studies, each of which may be fairly well designed internally, but the results of which are clearly incompatible with each other. Some of this discrepancy in the findings is undoubtedly due to variations in the methods used. But how are such variations to be weighted? By an egalitarian counting of the numbers of studies on each side of the issue? This is quite often done by unimaginative reviewers, but surely it must be regarded as an act of abdication from the task of building something resembling a science.

The tenacious hold which inductivist mythology acquired over the research practice of psychologists led to the delusion that the question of methodological bias need be addressed only in the context of the individual research study. As long as each study was well designed their piling up one on top of the other would somehow result in a scientific discipline. This was rather like the economic doctrine of the Hidden Hand of the market mechanism. As long as each individual contribution was rational in terms of its specific local context, the general good would be automatically and progressively realized. Unfortunately, the history of many areas of psychological research leaves little ground for optimism on this score. Rationality on the level of the individual research study is not a sufficient, and I suspect, not even a necessary condition for the cumulative success of the research enterprise as a whole. What does appear to be necessary for such overall success is a general commitment to an adequate theoretical framework.

Now, if we consider the relationship of method and theory on the level of a whole research area, instead of an individual study, there appear to be two possibilities to begin with. The variability of methods in use might happen to be such that different types of theory have an equal chance of being confirmed, other things equal; or the methods in use are systematically biased to favour one type of theory over another. The former possibility is clearly not realized in practice because of the existence of massive institutional pressures in favour of the preferential use of a certain research methodology. In the case of psychology that preferred methodology is based on the use of certain statistical techniques, the requirements of which govern the design of experiments. These statistical techniques are no exception to the rule that methodologies are not theory neutral but tend to produce results that are fundamentally biased to favour theoretical interpretations of a particular type. But if this is so, then the general use of these techniques would indeed constitute a kind of Hidden Hand that steers the research process as a whole in a certain direction. Whether that direction is to be applauded or deplored is a question that can only be addressed after obtaining some insight into the nature of the theoretical bias introduced by the methodology in question.

2. NUMERICAL SYSTEMS FUNCTION AS THEORETICAL MODELS

In the period since the Second World War psychological methodology has come to be dominated by the use of inferential statistics. There is no doubt about the *practical* usefulness of these techniques in diverse settings where clear decisions are required about the advisability of specific courses of action that involve known risk factors. However, 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. This assumption is associated with an implicit belief in the theory-neutrality of the techniques employed.

The imperative need to employ a limited number of techniques of statistical inference dominates the design of psychological experiments and even works back to the construction of data gathering techniques that must produce information of a type that can be handled by the statistical techniques. 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.

Very damaging, and essentially unanswered, criticisms of various aspects of the use of statistical inference in psychology have been in circulation for some time,[3] though they do not seem to have diminished psychologists' appetites for this methodology. In the present context I will not be concerned with most of these criticisms, but will restrict myself to the consequences that arise as a result of the institutionalized attempt to apply techniques of statistical inference to questions of systematic psychological theory. These consequences flow from the fact that the methodology of statistical inference is not a passive instrument but imposes a definite theoretical model which may or may not agree with the model offered by the theory being investigated by means of this methodology.

In exploring the reasons why statistical inference necessarily imposes its own theoretical model on research we must first note that the statistical techniques require numerical data for their application. Such numerical data can only be obtained through the imposition of a numerical system on some data source. It must be emphasized that this involves much more than a simple numerical labelling of empirical items—it means that the structure of the numerical system is taken to *represent* the structure of the empirical system.[4] Thus when engaged in research employing these methods one only knows the structure of the empirical system through the numerical system that forms the basis for statistical inference.[5]

There are two consequences for psychological theorizing, depending on whether one is using the methodology to develop theories inductively or to test already existing theories. In the first case, the structure of the numerical system will automatically be reflected in the theory. In other words, there will be a relatively straightforward determination of theory by methodology. On the other hand, if a theory is being tested that was formulated independently of the methodology there is a problem. In order to establish the relevance of the results obtained for the theory being tested one ought to be able to show that the structure which one's numerical system has imposed on the data is at least broadly congruent with the structure suggested by the theory. If it turns out that the numerical structure and the theoretical structure involve different assumptions, then the theory one is testing is not the theory one wanted to test but at best some vague analogue thereof.

The application of a numerical system involves a structuring of the domain into elements with certain properties and relations with certain

3 D. E. Morrison, and R. E. Henkel (eds.), *The Significance Test Controversy*, Chicago 1970.

4 P. Suppes and F. L. Zinnes, 'Basic Measurement Theory' in R. D. Luce, R. R. Bush and E. Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 1, New York 1963.

5 Of course in practice it is difficult to be rigorously consistent about this. See H. G. Petrie, 'A Dogma of Operationalism in the Social Sciences', *Philosophy of the Social Sciences*, 1, 1971, 145-60.

properties. The elements must be independently identifiable, e.g., specific responses on a psychological test. They must be such that one can form sets of two or more elements all linked by the same relation. Only a very few relations like equivalence or simple ordering make the application of numerical systems practically useful. These relations must have properties like transitivity (if $a = b$ and $b = c$, then $a = c$) in order to permit a valid application of numerical systems to an empirical domain. In other words, such an application involves a very definite theoretical structuring of the world one is interested in.[6] If the methodology is considered to be the *sine qua non* of scientificity, as it usually is, then there will be enormous pressures for the structure of *all* theories to accommodate to the theoretical structure embedded in the methodology.

Moreover, this embedded theory involves its own value hierarchy which adds to the basic constraints that it imposes on theorizing. Applications of numerical systems are not all on the same level but are graded according to the type of measurement scale which they enable one to construct. Four grades are commonly distinguished, involving nominal, ordinal, interval and ratio scales.[7] This categorization itself forms an ordinal scale, with 'ratio' at the top. The reason for this asymmetry is that the range of permissible arithmetic transformations and hence inferential statistical techniques increases as one goes up the scale of scales. Now if one's methodology is dominated by the requirements of statistical inference, it is all too easy to become enthralled by the more splendid vistas opened up by the additional techniques that become available to one as one improves the status of one's measurement scale. This methodology therefore contains inherent pressures to improve one's scale type.

But this has further theoretical implications because the higher the level of the scale the more structural assumptions must be made to allow the numerical system to operate. These additional assumptions impose further restrictions on the structure of the data one allows oneself to work with. Again, the desire to satisfy the demands of methodology leads to pressures on theoretical formulations to conform to the required structure. When the methodology is applied to already existing theories, the demonstration that the theory would lead one to expect this particular structure in the relevant observations is virtually never given. Moreover, the fact that this structure may be inherently implausible on the basis of the theory is usually ignored. The net result of such practices can only be the transformation of the original theory into something else.

6 G. Gigerenzer, *Messung und Modellbildung in der Psychologie*, Munich 1981.

7 There are a number of variants but the basic distinction goes back to S. S. Stevens, 'On the Theory of Scales of Measurement', *Science*, 103, 1946, 677-80.

3. STATISTICAL INFERENCE AND PSYCHOLOGICAL THEORY

In addition to the theoretical requirements that emanate from the necessity of imposing numerical systems there are other pressures that emanate from the statistical inference process itself. That process involves the use of samples in order to draw reasonably based inferences about hypothetical populations. The nature of the population one wants to hypothesize about will determine the sample that one investigates. In principle, there is no reason why one should restrict onself to sampling individuals; one could equally well sample situations or the actions of a single individual. However, there have only been rare instances of this in the psychological literature, and the overwhleming majority of published studies has always involved a sampling of individuals. There are both historical and philosophical reasons for this. Historically, the tradition of a methodology based on statistical inference was established at a time when most psychologists were extremely interested in the practical utility of drawing inferences about large populations of individuals and very little interested in individual cases. Philosophically, most psychologists seem to continue to believe that *aggregate* data from many individuals form the only acceptable basis for making *general* theoretical claims about individuals.[8] Hence the use of statistical inference in psychological research generally involves the use of group data obtained from a collection of individuals.

Now, this practice has some rather far reaching theoretical implications. In general, it is associated with two kinds of expectation. On the one hand, it leads to the assumption that the structure of psychological processes in individuals is isomorphic, or at least essentially comparable to the structure of group data. On the other hand, it is associated with the expectation that psychological theories should be such as to permit deductions about aggregate observations in samples of individuals. When applied in the context of induction these expectations impose clear limitations on the theoretical hypotheses that will be seriously entertained. They tend to be limited models that satisfy these two conditions. When applied to theories generated outside the requirements of the methodology the fact that these expectations are not fulfilled is generally ignored and the theories in question are simply assimilated to the prevailing standard of what a theory should be.

Psychological theories generated outside the now standard methodology were not generally in the business of making predictions to aggregate data obtained by sampling populations. Yet, although they avoided statistical inference, the work of men like Wundt, Freud, Köhler, Wertheimer, Lewin, and Piaget seems to have contributed in rather important ways to the progress of modern psychology. Their various alternative philosophies of science and of scientific method cannot

8 See D. Bakan, *On Method*, San Francisco 1967, chap. 1.

be analyzed here, but it is worth noting that at least one of them explicitly addressed himself to the question of the relevance of observed statistical regularities for psychological theory. Basing himself closely on the philosophy of Ernst Cassirer,[9] Kurt Lewin argued that the reference of scientific theory was the lawfulness of the genotypical level of events and not the fluctuating phenotypical conjunctions through which this lawfulness manifested itself in the empirical world.[10] The merits of Lewin's view are not at issue here. What is clear, however, is that no method which is based on theoretical principles directly contrary to those of Lewin could claim to be neutral with respect to his theory.

But Lewin's case is unusual only with respect to the explicitness with which the theoretical significance of overt statistical regularities is discussed. The disjunction between the aims of a theoretical system and the theoretical assumptions of statistical methodology is not peculiar to Lewin's field theory. Neither Piagetian psychology nor Gestalt psychology, and certainly not psychoanalysis or phenomenological psychology, would or could endorse the assumptions on which statistical theory testing methodology is based. The act of subjecting such theories to this kind of testing already presupposes that the theories are mistaken in some of their basic assumptions. All these theories therefore enter the test situation with an absurd handicap.

Leaving aside questions of psychotechnology, the use of the statistical methodology has been justified on two sorts of grounds—that it yields results which are more representative, and that it permits greater precision. The first of these hardly merits serious consideration. Given the practical realities of statistical sampling in psychology, it is difficult to see by what logic the responses of thirty American sophomores to the conditions of a psychological study, chosen from the infinite set of such conditions, might be considered more 'representative' than the responses of a single Viennese neurotic lying on Freud's couch. At best one might be able to claim that thirty American sophomores are more representative of all American sophomores (at a particular historical moment) than one Viennese neurotic was representative of all Viennese neurotics. But this would be both trivial and psychologically irrelevant.

The justification of statistical methodology on grounds of greater precision has some merit, but it is also necessary to recognize that the particular form of 'precision' advocated has enormous implications on the level of psychological theory. As we have seen, the production of observations of a form which makes them suitable raw material for the process of statistical inference involves the imposition of a certain

9 E. Cassirer, *Substance and Function*, New York 1953 (orig. 1910).

10 K. Lewin, 'The Conflict between Aristotelian and Galilean Modes of Thought in Contemporary Psychology', in *Dynamic Theory of Personality*, New York 1935 (orig. 1931). Also 'Gesetz und Experiment in der Psychologie', *Kurt-Lewin-Werkausgabe*, Bern and Stuttgart 1981 (orig. 1927), vol. 1.

structure on the data. Of course, observations must always be structured before they can enter into theoretical discourse. But the structures demanded by the methodology of statistical inference are of a particular kind, very different, for example, from the structures of natural language or alternative mathematical structures. The effect of this is to produce strong pressures to duplicate these structures on the theoretical level. For it is this condition which makes it a relatively straightforward matter to produce the required statistical predictions. As long as the theoretical model is isomorphic with the empirical model that defines the pattern of observations, specific elements of the one can be co-ordinated with specific elements of the other in a fairly unambiguous manner. Under these conditions statistical prediction from the theoretical model to the empirical model and statistical inference from the empirical model to the theoretical model is relatively unproblematic. But if the two levels are not isomorphic these operations cannot be carried out without mediating transformations that will involve additional assumptions and of course a change in one or both of the models. In the specific case under discussion here the structure of the empirical model is fixed by an institutionalized methodological imperative. Hence it is the theoretical model that will undergo transformation.

There are two considerations that must raise very serious doubts about the wisdom of allowing this process to take its course. In the first place, the criteria that determine the structure of the empirical model are logical and mathematical criteria that are psychologically irrelevant. With the transformation of the theoretical model to accord with the structure of the empirical model one is apt to end up with a representation of psychological processes that treats them as though they were logical processes. This is undoubtedly a major source of certain forms of 'cognitivist' bias in contemporary theorizing.

The second consideration involves a more pervasive problem. The most remarkable aspect of the application of statistical inference to questions of psychological theory is that in the usual case the empirical model and the theoretical model refer to completely different levels of reality. The empirical model refers to the structure of an aggregate of observations gathered from a *group* of individuals; the theoretical model refers to processes associated with a single *individual*. There certainly are no *a priori* reasons why we should assume *any* structural similarity between psychological processes in an individual and the logical framework which we have imposed on our group data. In fact, such an assumption appears to be inherently implausible. These are totally different orders of events—why should we expect them to be structurally comparable? The answer is of course that they are not comparable until psychologists have made them so by constructing theoretical models of psychological processes in the image of the logical structure of their group data.

The first attempt in this direction was Thorndike's 'connectionist' model of the mind as an aggregate of independent elements varying in strength over time.[11] If one regarded group data as constituting a sampling of these elements one could make appropriate inferences from the one level to the other. Variants of this basic model continued in use among learning theorists for a long time.

But the most common type of theoretical model designed to bridge the chasm between psychological process and statistical aggregates is more complex. The imposition of numerical systems on group data, that was referred to earlier, results in a grouping of such data in terms of specific scales or variables. This is a precondition for the application of the commonest forms of statistical inference. The formal properties of the numerical systems used in their construction are reflected in the nature of the relationships between these ordered sets of observations. Only purely formal relationships like addition, multiplication, distance etc. can be accommodated by these models. Moreover, for these relations to operate it is also necessary for each variable to be defined independently of any other and to remain identical with itself irrespective of changing circumstances. In other words, intrinsic relations and qualitative changes are excluded. They could only enter the system by undergoing an appropriate transformation.

As long as these principles are simply applied as rules for producing a certain kind of empirical order, the only objections one could raise would be governed by practical considerations. However, it is common practice in psychology to duplicate these empirical structures on the theoretical level so as to permit easy passage from the system of group data to the psychological system of the individual. This results in an account of psychological structure in terms of rigid constructs linked by purely formal relations. The term 'operational definition' is generally used to refer to this duplication of the empirical structure of group data on the level of individual psychology. Of course, the crucial 'operation' here involves the psychological reification of structures originally imposed on data because of methodological requirements. (The term *methodomorphic* theory seems to provide a shorthand description of this state of affairs). In this way the tricky problem of the passage from statistical data to individual psychology and back again is solved, but only at the cost of letting special methodological requirements limit the structure of one's theoretical concepts.

These practices commonly lead to the forced assimilation of heterogeneous theoretical content to the model that is implicit in the methodology. Whenever theories that do not share the theoretical assumptions of standard methodology are subjected to examination by

11 E. L. Thorndike, 'A Note on the Accuracy of Discrimination of Weights and Lengths', *Psychological Review*, 16, 1909, 340-46.

means of that methodology they must first be translated and reformulated in terms that the methodology can handle. But in this process many of the characteristic features of these theories are lost, and they are transformed into something quite alien to their original conception. They are apt to end up as just another variant of the standard methodomorphic form of theory. Insofar as they are then taken seriously only in this changed form, the effective range of theoretical options is further impoverished.

In this way the implicit assumptions of the methodology come to function very much like a set of axiomatic principles that defines the limits of acceptable research and theorizing. It is these methodologically embedded principles, rather than any explicitly formulated theoretical propositions, that function as the unifying concepts for the discipline, or at least for most of it. The conspicuous lack of unifying theoretical concepts in the discipline of psychology has often been noted.[12] Yet, a degree of coherence does exist in spite of this. My suggestion is that a major source of such coherence lies in a widely accepted methodology that in fact involves certain basic theoretical commitments. It is this shared commitment that provides the basis for effective intra-disciplinary communication. The gradual emergence of the methodological imperative appears to have been the dominant theme in the development of twentieth century psychology.

4. CONDITIONS OF PROGRESS

Not the least important consequence of the operation of the methodological imperative is that it defines the conditions of progress for the discipline. It is clear that the theoretical assumptions which are built into the standard methodology can never be refuted by the use of this methodology. The fundamental features of the data base are predetermined by the methodology that produces it. These are really stipulated rather than discovered features. As long as the methodology enjoys overwhelming social support within the scientific community these features are protected from the effects of contrary evidence because no evidence from outside the charmed methodological circle is accepted as valid. Within the circle no anomalies can arise, at least not with respect to the forms that the method necessarily impresses on the evidence. Theoretical change can occur only within the limits prescribed by these forms. A more fundamental theoretical change would surely depend on a fundamental change in methodology. It is a mistake to believe that any specific set of rules about what is to count as scientific evidence can be fixed independently of the content of achieved scientific theories and of the goals of research. On the contrary, such 'rules of evidence' in

12 S. Koch, 'Language Communities, Search Cells, and the Psychological Studies', in W. J. Arnold (ed.), *Nebraska Symposium on Motivation*, Nebraska 1975.

science have always been subject to change in association with changes in content.[13] The history of twentieth century psychology itself provides striking examples of this. It is when a field takes a single set of narrowly conceived techniques as the sole index of scientificity that it is likely to condemn itself to the impotence of rituals.

Of course, it may be held that the established methodological framework in psychology provides ample scope for the development of a healthy variety of theoretical ideas. But the historical evidence is not kind to this point of view. The reigning methodology has been established long enough to permit an informed appraisal of its promises and failures. There is no question that it can be very effective in the context of specific practical problems that involve questions of limited scope which require an unambiguous answer. These are problem solving situations with a rather similar structure to those in agriculture and industry in which the basic approach of the standard methodology of statistical design and inference was first developed.[14] In this kind of context the assumptions of the methodology have a certain practical utility. Pragmatic theories developed in connection with this type of situation also lend themselves readily to testing by means of the same methodology. An extension to other theories with essentially pragmatic goals is also possible without raising too many problems. In fact, it is this rather unambitious kind of theorizing that has flourished under the protection of the traditional methodology.

What results from this procedure is a collection of generalizations that describe relations among classes of variables, the kinds of relations and the kinds of variables being predetermined by the methodology. If one is content to accept that this is sufficient to establish the scientific credentials of the field, then there is no cause for further questioning. Progress is then simply a matter of a quantitative growth in data and in generalizations from these data, always of the same type. It is only if one tries to go beyond these modest aims that problems arise. If one begins to look for systematic theoretical coherence in terms of generative processes that function as explanations for observed regularities one has a choice. One possibility is to limit oneself to developing theoretical models whose properties are congruent with the properties that methodology has imposed on the data. In this case the priority of methodology remains undisturbed.

Another possibility is to work with models, developed outside the influence of the methodology, which make assumptions that are at

13 D. Shapere, 'The Character of Scientific Change'. *Boston Studies in the Philosophy of Science*, Boston 1980, vol. 56.

14 See the classical texts of R. A. Fisher, *The Design of Experiments*, Edinburgh 1935, and G. W. Snedecor, *Statistical Methods*, Iowa, 1957, chap. 14. Also L. T. Hogben, *Statistical Theory*, London 1957, chap. 14; and D. A. MacKenzie, *Statistics in Britain*, London 1981, pp. 111-16.

variance with it. Examples of such assumptions are structuralist assumptions, assumptions about epigenetic transformation, about overdetermination as a necessary feature of human functioning, about intrinsic relations within symbolic content, about non-additive system properties, etc. In these cases a test of the theoretical model by means of the standard methodology will always be singularly unconvincing. For what is tested is not the theoretical model but some transformation of it to suit the requirements of the method. Hence both positive and negative outcomes are equally irrelevant. An examination of attempts to subject Freudian notions to a test by conventional experimental designs suggests that both friends and foes of these views seem to have little difficulty in drawing sustenance from such tests.[15] It would reduce conceptual confusion to make a clear distinction between the classical versions of these theories and their transmutations through methodological alchemy.

At the most general level one could characterize the methodomorphic style of theorizing as the subjection of psychology to logic, or allowing logical categories to function as psychological categories. Historically this tendency can be seen as part of the heritage of classical empiricist philosophy. The tendency to confuse logical and psychological issues was highly characteristic of this school of thought, but because its critics were generally philosophers their criticism usually took the form of objections to the psychologizing of philosophical issues. However, this coin has another side, as was noticed by Wundt, the founder of psychology as an experimental discipline. Wundt's criticism of his predecessors was that they had substituted a 'logical schematism' for genuinely psychological principles.[16] He recognized that the establishment of his new psychology would involve a constant struggle against the temptation to substitute logical for psychological categories.

In his own personalizing style William James, the rival pope of the new psychology, expressed something similar with his notion of 'the Psychologist's Fallacy', This referred to the psychologist's tendency to substitute his own categories for those of the person being studied.[17] Both Wundt and James were reacting against Mill's *Logic*, that ultimate celebration of the union of logic and psychology which leaves the latter with the status of a deductive science. However, by the middle of the twentieth century Mill was back in favour among psychological methodologists with a sense of history.[18] Embedded in a sacrosanct

15 Cf. H. J. Eysenck and G. D. Wilson, *The Experimental Study of Freudian Theories*, London 1973; and S. Fisher and R. P. Greenberg, *The Scientific Credibility of Freud's Theories and Therapy*, New York 1977.

16 W. Wundt, *Grundzüge der physiologischen Psychologie*, Leipzig 1887, 3rd ed., vol. 2, chap. 17.

17 W. James, *The Principles of Psychology*, New York 1890, vol. 1, p. 196.

18 E. G. Boring, 'The Nature and History of Experimental Control', *The American Journal of Psychology*, 67, 1954, 573-89.

institutionalized methodology logic had returned to claim its ancient dominion over psychology.

Of course, a plea for the rejection of this domination does not entail the abandonment of disciplined experimentation or methodological rigour. This was certainly not on Wundt's agenda, and it is not on the agenda now. The issue is one of the relation between psychological theory and the rules of evidence. Three commonly held beliefs affecting this issue appear to be ripe for revision: (1) that statistical inference provides the only valid procedure for relating data and theory; (2) that the rules about what constitutes valid evidence are independent of theory and are fixed forever; (3) that the structure of theory must be accommodated to the structure of methodology and not vice versa. As long as these views prevail theory testing in psychology will be a matter of choosing among different versions of a theoretical position, the fundamental features of which are in fact beyond dispute.