Chapter 10
Statistical Models, Scientific Method and Psychosocial Research

Raymond J. Adams

Abstract This piece is a compilation of a number of short class-notes I wrote in 1987 and 1988 as a result of discussions with Ben and fellow students whilst I was a student at the University of Chicago. At that time Ben was pushing us to consider why progress in the psychosocial sciences seemed to be so frustratingly meagre when compared to progress in the 'hard' sciences. In discussions with Ben it seemed, to me at least, that central to his argument was a view that much of 'so-called' statistical modelling was unscientific—that it focussed on the description of ad-hoc collections of existing data, rather than proposing and rigorously testing of models and theories through the analysis of measures with well understood properties. Ben was very critical of exploratory statistical analysis, made a clear distinction between measurement models and analytic models and was always reluctant to fit statistical models to data—he wanted to use statistics as a tool to test whether data were consistent with theoretically posited models, he wanted to fit data to models. One wonders how he would have felt about the current big data and data mining movements.

10.1 Note

The material below is not meant to be profound, nor should it be read as presenting a well-developed view on statistical modelling. What I hope it does is give an insight into the nature of the discussions Ben held with his students and in class during those days.

10.2 Psychosocial Research Methods

Psychosocial research involves the study of human behaviour and social phenomena. Depending upon the context it has been referred to as 'social science,' 'human science,' 'social inquiry' or 'behavioural science.' These labels are

R.J. Adams (✉)
ACER, Melbourne, VIC, Australia
e-mail: ray.adams@acer.edu.au

M. Wilson, W.P. Fisher, Jr. (eds.), *Psychological and Social Measurement*, Springer Series in Measurement Science and Technology,
https://doi.org/10.1007/978-3-319-67304-2_10

generally meant to encompass a core of disciplines such as Sociology, Psychology and Economics, and fields of application such as Education and Social Welfare.

The role of scientific method, as an approach to understanding the human world, has become an ongoing issue of debate in psychosocial research. Indeed, the philosophy of science is filled with debate about the validity of scientific method (in any context) (see for example Chalmers, 1976; Polkinghorne, 1983; Phillips, 1987). Despite the debate, it is recognised that scientific method has been successful in its application to the study of the natural world and, recognizing this success, workers in psychosocial research have attempted to apply scientific method to the human world.

In this paper, I do not intend to produce long and involved arguments regarding the definition of scientific method, as it is applied to psychosocial research, it will be useful for later discussion to present a simple framework for describing science. The framework that is presented is a personal one but draws upon Fowler (1962), Chalmers (1976), Polkinghorne (1983) and Phillips (1987).

I will then mention some of the problems that have been encountered in the application of scientific methods to the human world and in the final section I will discuss the application of statistical models to psychosocial research. The application of these methods may play a role in the disappointing outcomes of scientific enquiry in psychosocial research.

10.3 What Is Science?

Science is the dominant process employed by man to achieve a knowledge of the world. Science is driven by a view that knowledge can be expressed as a set of publicly accepted and infallible truths. Science never claims to have actually found these truths but it does claim to be getting closer to them. This identifies an important fallacy regarding science that should be dispelled. “Science” does not mean “truth.” Hence, the term “scientifically proven” as is claimed by so many television commercials, is a logically inconsistent statement.

Science, like all desires to understand the world, is motivated by a need to survive. All areas of man’s inquiry develop from investigations originated to increase chances of survival. In its motivation, science is no different than any other method proposed as a tool to help us understand our world.

There are no definitive characteristics of science but it can be distinguished from other methods of inquiry by the principles of knowledge on which it is built. The first is the principle of objectivity, public truth and verifiability. For knowledge to be scientific it must be possible for that knowledge to be independently verified. At this point we are walking a fine line, since “objectivity, public truth and verifiability” could easily be construed as requiring absolute truth.

This is not however the case–science does not require knowledge to be absolute; what it requires is the possibility of verification and publication, within the constraints of a given, shared set of values and a common theoretical stance. A second principle is that the outcomes of science must be useful. This does not mean that science must always fulfil a specified need, often the practical uses of scientific discoveries are not identified until well after a discovery has been made. In fact, there is a strong argument to be made for science as a creator rather than satisfier of needs. Science is useful when it leads to a simpler yet more comprehensive understanding of the world. Eisner (1979) combines the principles of objectivity as follows.

> … objectivity is a function of intersubjective agreement among a community of believers. What we can productively ask of a set of ideas is not whether it is really true but whether it is useful, whether it allows one to do one's work more effectively. (Eisner, 1979, p. 214)

The third principle for science is that our current knowledge is incomplete and represents only a construction that attempts to explain reality and that construction must always be open to development, modification or rejection.

The process of science is one of making observations and making inductions from those observations to develop a "theory." The method of observation must always be consistent with the principle of objectivity and the process must be ongoing in recognition of the principle of incomplete knowledge. Kinston (1985) spells out five stages that are involved in scientific work. I will take some time to describe each of Kinston's stages since they are important in the following discussion.

Kinston's first stage, level I, is "entity." All knowledge begins with the formation of ideas. In examining reality, the scientist begins by creating an idea or concept. These concepts can be very general. For example, they could be anything from "ability" to "heat" or even quantity. The entity is subjectively defined.

Level II, "observable," involves two ideas. We must take our original idea and add "thingness" so that our original idea or concept can become public. In practice, Level II involves the definitions of conditions and criteria that enable the original concept to be operationalized.

At Level III, "comparable," the concept of quantity is added to "thingness" and our original idea. According to Kinston "a comparable is formed by ordering and ranking observables and answers the question: 'which is more (less)' or 'which is better (worse)?'" (p. 98). According to Kinston, Level III requires the subjective use of the concept of quality.

For Level IV, "measurable," we add the idea of "generally applicable unit." Quantity is taken for granted and we establish a measurable that enables us to describe "how much" in an absolute sense. Clearly, that absolute must be defined with respect to some standard. Kinston sees the move from levels I to II and from III to IV as a process of objectifying the idea. This use of objectivity corresponds to the one used as a principle of science. That is the "idea" is becoming more public, within a specified set of constraints and conditions. Note that subjectivity also plays a role in science. This subjectivity plays a role at level I, where the original idea is private, and at level III where the subjective sense of quantity is introduced. We attempt to impose objectivity by moving from levels I to II, and from levels III to IV.

The last of Kinston's levels is Level V, "relatable." The idea of "relation" is added to our previous four and we begin to describe the world in terms of relationships between the subjective ideas we began with.

According to Kinston then, the result of scientific inquiry is the development of "theories" that specify relationships between measurables in "a deliberate attempt to model or represent significant aspects of reality" (p. 95). We could add that these theories will support predictions and explanations of events within a defined class. The boundaries of that class are specified by the contents and specifications used in the development of the entities, observables, comparables, measurables and relatables that make up the theory. The principle of incomplete knowledge warns that scientific theories are not infallible and they must always be open to modification. The scientist must always be prepared to go back and modify the construction at any level to improve the utility of the theory.

While there is no way for scientific theories to be proven correct they must always be open to refutation. In fact, Popper sees testability and openness to refutation as the essence of scientific inquiry (Phillips, 1987).

10.4 Science in Psychosocial Research

The effectiveness of scientific methods of inquiry in the natural sciences has led to its adoption as a paradigm for psychosocial research. But even the most ardent proponents of scientific methods in the human sciences have recognized that the achievements thus far have been a little disappointing. For example, Hedges (1987) comments that:

> Psychologists and other social scientists have often compared their fields to the natural (the "hard") sciences with a tinge of dismay. Those of us in the social and behavioral sciences know intuitively that there is something "softer" and less cumulative about our research results that about those in the physical sciences' (Hedges, 1987, p. 443)

Many factors have been identified as possible causes for the apparent failure of scientific methodology in the psychosocial sciences—Hedges (1987) lists a number of references that discuss possible explanations for the perceived failure and limitations of scientific methodology in psychosocial research. Valentine (1982) emphasizes two possibilities. The first is a lack of systematicity. She believes that science relies on systematicity in the subject matter so that a coherent body of knowledge can be developed and that a lack of systematicity in the human world causes problems in the definitions of variables that are suitable for the expression in a coherent body of knowledge. The second is generality. She claims that scientific theories are unrestricted by space and time, an ideal that cannot be met in research on the human world. The arguments against the suitability of scientific method in psychosocial research can be persuasive and many are not without merit.

The application of scientific method to the human world may well be more difficult than the application of scientific method to the natural world. But when identifying sources of failure for a particular research paradigm we should not only examine the subject matter and its suitability for us with the paradigm, but we should also examine the fidelity with which the paradigm was employed. Before we begin to criticize the appropriateness of the scientific method in psychosocial (because of its apparent failure) perhaps we should examine the fidelity with which scientific method has been employed.

Some of the most common misconceptions in psychosocial research have centered around the use of quantitative data, experiments and sophisticated statistical models. It would not be unfair to argue that most researchers believe that the more of these three factors you have, the more scientific your study is. Perhaps this is part of the problem of scientific method in psychosocial research. Quantitative data, experiments and statistical models do not make science.

In the remainder of this paper, I will examine the case of statistical models and comment on when their use may be scientific and when it may not.

10.5 Models

The term *model* has widespread use throughout all research. We have for example: The general linear model, models for pattern recognition, internal models of attachment figures, computer models of learning processes, stochastic models for learning and Rasch models, to name just a few. In each case a model acts as a representation of reality. Models have proven fundamental in all forms of inquiry and they have a central role in scientific method.

In most cases models are expressions of theories but the form of the expression will depend on the purpose of the models. First, a model can be used to assist in the *explanation of theory*. This is usually done by constructing the model with terms, concepts and images that are more readily understandable than the theory itself. An important purpose of models lies in the *testing of theory*. When formulated as a model, logical inconsistencies in the theory may be identified. In some form, models can be tested through simulations of reality and, when expressed in particular mathematical forms, a range of statistical methods are available to "test" the theory. Through improved explanation of theory and testing of theory the models can them lead for further development and enhancement of theory.

One of the largest classes of models used in psychosocial research are the statistical models—or the "off-the-shelf" variety of statistical methods. Note that those mathematical and statistical models that were developed for a specific purpose or research situation and that do not enter common usage are not meant to be covered by this discussion.

It is not an uncommon view amongst social scientists that the use of these models leads to a scientific research study. But to what extent do these methods act adequately as models? What role do these methods play in the scientific process described above? Based on these considerations, when are these models applied scientifically?

Table 10.1 A classification of statistical methods

Name	Example methods
Descriptive methods	Exploratory factor analysis
	Descriptive statistics
Explanatory methods	Log-linear modelling
	Stepwise regression
Confirmatory methods	Linear structural relations
	ANOVA
	Confirmatory factor analysis
Axiomatic methods	Rasch measurement

10.6 Categories of Statistical Models

To discuss their application to psychosocial research it is useful to construct a fourfold classification of statistical models. One possible classification scheme is presented in Table 10.1. The allocation of an approach to a category may depend on the mathematical form of the model, but it is more likely to depend on the methodological reasoning underlying the use of the method. For example, factor analysis, depending upon the details of its application may be classified as an exploratory or confirmatory approach. Each of the methods can be discussed in terms of the relative role of: the data, substantive theory and the constraints imposed by the mathematical form of the model.

10.6.1 Descriptive Methods

These are used to "fish around" in data. When the researcher has a body of data that has been observed and has no theory, it is possible to use mathematical methods to manipulate the data in a search for relationships that may be useful in a development of the theory. Perhaps the two most common exploratory methods, beyond simpler descriptive statistics are correlation and factor analysis. In many instances when a research (data analyst) is faced with a body of numerical data for which he/she has no theory, a set of correlations will be calculated to identify any covariation between variables. Substantive theory is then built to explain the observed covariation. Exploratory factor analysis is a more systematic approach to the examination of correlations. The aim of factor analysis is "the resolution of a set of variables linearly in terms of a small number of 'factors'" (Harman, 1976, p. 4).

In these techniques, the appropriateness of the model for the data is rarely considered. Supposedly, the statistical techniques employed allow the patterns and relationships in the data to be exposed, while making only very weak constraints in that exploration. In descriptive methods, it is hoped that the analysis is driven by the data, with theory playing only a limited role through the mathematical specifications of the model. In some of the simpler descriptive techniques such as scatter plots, histograms, box and whisker and the like, this assumption may be almost fulfilled. Beyond that things become less clear. Exploratory factor analysis and cluster analysis are obviously method bound and even the selection of a measure of central tendency (mean, median, mode) can have an impact on data interpretation.

10.6.2 Explanatory methods

These methods are used when developing models to describe a set of data. Their aim is to develop a mathematical model that accurately reproduces the observed data. Rather than being driven by theory these methods are driven by a combination of the data and the mathematical technology being employed. In general, the measure of success in applying these models is the degree to which the developed model confirms to the observed data (fit). In the development of these models there is always some tradeoff between model simplicity and the accuracy of the model in reproducing the data. The researcher must be careful to ensure that the plausibility, utility, elegance and simplicity of the associated theory does not get lost in the search for model to data fit.

In these explanatory methods, the form of the model places strong constraints on the development of substantive explanation. This is argued as valid on the basis that some models can generally be constructed to fit any given set of data. Unlike descriptive methods however, it is recognized that any structure identified, or developed from the data, is strongly bound by the researcher's approach to the analysis.

10.6.3 Confirmatory methods

Methods of this kind are used to test the plausibility of a theory when it is stated in a particular form. This category includes traditional approaches to experimental data analysis and the more recently developed confirmatory data analysis procedures. In both cases a mathematical model is constructed that is argued to be commensurate with the substantive theory to be tested. Mathematical and statistical techniques are then used to test the compatibility of the theory, as expressed by the model, with observed data. The most common approach in psych-social research is to take an 'off-the -shelf' statistical method and assume that it can be used to represent the substantive theory, then apply standard testing procedures designed for that method. The aim is to fit the model to the data. If the data does not fit the model, then the model is rejected and the theory (or the data collection method) is modified.

In the experimental case, the purpose is to test the plausibility of a specific hypothesis that has been proposed by the researcher. A mathematical model that is claimed to be commensurate with substantive theory is selected, and the model is then fitted to the data, and the acceptability of that fit is examined. In experimental designs, the model is formulated so that rejection of the ("null hypothesis") model adds support to the researcher's theory.

These approaches are driven more strongly by theory than descriptive or explanatory methods. Theory is used to construct the form of the models and the theory, as represented by the model, is tested through fit to observed data. While the data is not allowed to "speak for itself" in the sense of descriptive and explanatory models, it is being used to test the possibility of a particular theory being true.

10.6.4 Axiomatic Methods

The mathematical models used with these methods are derived from a set of axioms required by the researcher. It has been argued that in some instances these mathematical models are deduced from the axioms. By definition, these axioms cannot be proven or disproven.

Examples of axiomatic methods are the application of Rasch models. Rasch models are developed from fundamental axioms regarding the desired or necessary nature of measurement. Given these axioms it is argued that if a measuring instrument is to be valid in the sense of having specific objectivity, then it must conform to an appropriate Rasch model. Specific objectivity means that, once calibrated, the data from any subset of fitting items may be used to measure a person, and vice-versa, that the data from any subset of fitting persons may be used to calibrate the items. When developing the measuring instrument, observations are made and an attempt is made to fit the observations to the model. If the data do not fit the model, then it is argued that the instrument does not provide a valid ("specifically-objective") measure. The researcher should then examine why his/her measurement intentions have not been met by the instrument that was constructed.

In this case the mathematical form of the model, which has been built upon a set of specifications (axioms), takes a dominant role. Theory and model are far more intimately related than in any of the other approaches. The theory and model may in fact be the same thing only expressed in different forms.

10.7 When Are Statistical Methods, Models?

Harré (1976) considers two types of models: *sentential* models and *iconic* models. A sentential model is a set of sentences in some kind of correspondence with another set of sentences. An iconic model is a thing, structure or process in some kind of correspondence with another thing, structure or process. Harre further adds that models whose subject and source differ have come to be called *paramorphs* and those whose subject and source are the same are called *homeomorphs*.

Just as paramorphs may be the subject matter of sentential models, so too may homeomorphs. The description of homeomorph may be treated as a sentential model of the description of its source subject. I am inclined to think that this is the kind of modelling that we hope to do when applying statistical methods in psychosocial research. The sentences in the statistical method can be treated as a description of a homeomorph of the real psychosocial world.

The requirements of this kind of modelling include a correspondence between the sentences that make up the statistical model and the homeomorph that the researcher has constructed as a theory.

If we consider descriptive methods by these criteria we can see that they are not models at all. They are never intended to have any correspondence with a particular model of the psychosocial world.

Explanatory and confirmatory methods are both attempts at sentential descriptions of homeomorphic models and therefore do aspire to be legitimate models. Explanatory methods are attempts to build sentential models in the form of formulations of the relationships in observed data and confirmatory methods are both that and also attempts to test specific models against observed data. The validity of these models, for this purpose, depends upon their ability to reflect the researcher's homeomorph. Unfortunately, beyond the selection of variables for inclusion into the model this is rarely a major consideration of the practical researcher. The use of a method is often determined as much by its availability as its suitability for the problem at hand.

The axiomatic models belong more clearly to the class of sentential models. If we take the Rasch model as a particular example, then we can see that it forms a sentential model of the process of measurement.

In summary, we can see that descriptive methods are not models (and probably have no aspiration to be models). Explanatory and confirmatory methods need to be models if their application is to be valid; but they too often fall short in their correspondence with the researcher's other expressions of the same model. Finally, axiomatic methods are always models because they are built to be representations of specific theories.

10.8 What Role Do These Methods Play in Science?

The role of these methods can be further examined by looking at their place in the process of science as outlined by Kinston and discussed above.

All of the methods assume at least the first two of Kinston's levels—'entity' and 'observable.' Since all of the methods require the use of observations, these two levels must be first developed by the researcher. Although entity and observable must be defined before the application of statistical methods, one of their important uses is the provision of information for the modification and redevelopment of entities and observables.

In considering our four statistical methods it is our example of an axiomatic method, the Rasch model, that plays a unique role. The Rasch model is concerned with taking developments from the first three levels and constructing 'measurables' whereas descriptive, confirmatory and explanatory methods take observables and 'measurables' to produce 'relatables'.

Some explanatory, data fitting, methods are used in an attempt to construct measurables from comparables. The application of these methods to constructing measurables is however invalid. Since measurables can only be constructed by the addition of the entity 'generally applicable unit' to the comparable, the only valid statistical method is one that can govern the construction of that unit. An axiomatic model that is built specifically to take comparables and add generally applicable unit to provide us with measurables is necessary. The explanatory methods used for this purpose are not commensurate with the entity 'generally applicable unit' so they cannot be used to construct a measurable or test it.

While the range of descriptive, confirmatory and explanatory can be applied with data from the 'observable' or measurable levels, they are at their most powerful when they take measurables and examine the relationships between them to provide relatables. In some cases however, the nature of what we are studying may force us to search for relatables with observables or comparables.

10.9 When Are These Methods Applied Scientifically?

Each of these methods, if used wisely, has something to offer science although some are more likely to be of use than others. The Rasch model is a fundamental tool in making the construction of measurables possible and measurables are the most powerful variable that we can use in producing relatables. The confirmatory methods, if used with variables from the highest level feasible, and, when designed to be commensurate with substantive theory, are a powerful means of testing theory. But, if we do not ensure that the model matches the theory and the variables we use are of the highest possible level, then as Kinston (1985) warns "Plausible, satisfying and apparently meaningful fantasy may result" (p. 101).

Explanatory methods, even when used with measurable, are totally constrained by the selection of an arbitrary statistical procedure and they play upon possibly incidental patterns in the data. That is, in their attempt to identify the common variance between variables, they are in danger of taking positive advantage of noise and random elements of the data.

The descriptive procedures play mixed roles in science. Graphical methods like scatter plots, histograms and box and whisker plots can be useful tools in many aspects of data analysis. However, the other more "complex" procedures that are used are method-bound and in many cases, arbitrary in their findings leading to theory conflation and confusion.

10.10 Concluding Comment

It is interesting to re-read and reflect on this piece some 30 years after it was written and just under 28 years since my last exchanges on this with Ben. Whilst is has a certain naivete it is possible to see in it threads that have had a profound influence *on how I have approached a career of research and development work. But, in addition, it is with a tinge of disappointment that I note how so many of Ben's observations concerning the unscientific nature of so much so-called statistical modelling remain true today. Moreover, it is unclear whether the rapid recent development in machine learning will lead to an eclipse of this confirmatory approach to science, and to a domination of exploratory methods, especially given the breathtaking expansion of what we now consider to constitute "data."*

References

Chalmers, A. (1976). *What is this thing called science?* Brisbane: Queensland University Press.

Eisner, E. W. (1979). *The educational imagination: On the design and evaluation of educational programs*. New York: Macmillan Publishing.

Fowler, W. S. (1962). *The development of scientific method*. New York: Pergamon Press.

Harman, H. (1976). *Modern factor analysis* (3rd Edition Revised). Chicago: University of Chicago Press.

Harré, R. (1976). The constructive role of models. In L. Collins (Ed.), *The use of models in the social sciences* (pp. 16–43). Oxford: Tavistock Press.

Hedges, L. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. *American Psychologist, 42*, 443–455.

Kinston, W. (1985). Measurement and the structure of scientific analysis. *Systems Research and Behavioral Science, 2*, 95–104.

Phillips, D. C. (1987). *Philosophy, science, and social inquiry*. Oxford: Pergamon Press.

Polkinghorne, D. (1983). *Methodology for the human sciences*. Albany: State University of New York Press.

Valentine, E. R. (1982). *Conceptual issues in psychology*. London: Allen and Unwin.