1 The Experimental Tradition

GLENN REGEHR
University of Toronto

SUMMARY

This chapter examines the experimental tradition in educational research. With it, an attempt is made to provide a sense of the background assumptions and everyday details that are involved in the process of experimental research. Of course, a chapter cannot be comprehensive, and this chapter is not intended to be prescriptive. Rather it is intended to be a guide, an approach to thinking about experimental paradigms. Thus, issues and examples are selected that are particularly relevant and helpful in the role of an experimental researcher and a teacher.

The chapter is divided into several sections. The first section will spend some time providing a historical and epistemological framework for understanding the tradition of experimental research. The second section will define experimental research in its formal sense, but will also provide some feel for the ways in which the formal version of experimental research has been adapted and the implications of those adaptations for altering and limiting the conclusions that can be drawn from the research. The third section pays more attention to the specific steps associated with developing an experimental research study. It will provide a perspective on the development and refinement of a research question, discuss the details of selecting (or creating) relevant outcome measures and conceptualizing appropriate comparison groups (control conditions). Then a discussion of the original question will be re-addressed to reflect on its evolution as we have been working on the details. The final section will focus on the strengths and weaknesses of some specific research designs which are commonly used in education.

Before beginning, however, a note on the philosophy of the chapter. The concept of an unbiased opinion is an oxymoron. This is as true in the domain of experimental research as it is anywhere else. Biases are inherent in every aspect of the experimental process including: what question is asked, the way the question is

International Handbook of Research in Medical Education, 5–44.
G.R. Norman, C.P.M. Van der Vleuten, D.I. Newble (eds.)

asked, the method of research used, the types of comparisons made, and the interpretations that arise. The point of this chapter is not to provide efforts to eliminate those biases. This is impossible and therefore a fool's errand. Rather, the intent is to highlight some of the biases, to make them explicit, and to provide some framework for helping the researcher make his or her own decisions about what biases will be acceptable for the intended purpose. Similarly, it is important to realize that no experiment is ideal. There are compromises inherent in every design. Thus, this chapter is not a recipe for developing the "right" research design. Rather the discussion will describe the strengths and weaknesses of each design and leave it as an exercise for the reader to determine which combination of strengths and weakness is optimal for his or her current purposes. This perspective is not meant to sound cynical or to be discouraging. There is no doubt that experimental research is a powerful and useful tool for expanding and refining knowledge. But like any powerful tool, it must be used with caution, respect, and awareness. It is with this perspective in mind that the following sections are written.

A HISTORICAL AND THEORETICAL CONTEXT

The tradition of experimental research has a long history. As early as the seventeenth century scientists such as Descartes (1596-1650) were using the "scientific method" to discover the "laws of nature". Inherent in the desire to discover these "laws" was a drive to understand how nature worked, the underlying mechanisms that caused events to occur. Thus, the paradigm of experimental research was developed to provide grounds for claims of causality.

The discussion of causality in this context must be considered at two levels. These scientific pioneers were attempting to create comprehensive *theories* regarding the workings of natural phenomena. That is, they were trying to establish the *underlying* causes of events, the reasons *why* a particular phenomenon occurred in the way it did and when it did. To do this, it was necessary for them to get control over the phenomenon of interest, to identify functional variables that caused an event to occur. The theory provided a set of systematic, lawlike generalizations that predicted the nature of the causal relationship that ought to exist between a particular variable and the phenomenon of interest. Thus the experiments were physical extensions of the theory in the form of specific hypotheses. By testing the hypothesis the researchers were empirically testing the theory, confirming, extending and/or limiting its applicability and generalizability. In turn, the theory provided a structured, systematic framework for predicting the scope and relative strength of the particular causal relationship seen, and provided a roadmap for determining the next logical step in the experimental program of research. Thus, causality at the level of the theory is abstract and thereby allows for generalization and prediction. Causality at the level of the experiment is functional and thereby allows for demonstration of practical control over the phenomenon of interest.

The use of the scientific method was formally applied to learning with the advent of psychology and education as fields of study (James, 1890; Thorndike, 1903). Here much of the emphasis was on the development of the set of universal laws that governed learning, with much of the research, at least in psychology, being carried out on animals other than humans (for example, Thorndike, 1898). In psychology, this theoretical paradigm thrived under the name *Behaviorism* (Watson, 1925; Skinner, 1938) until the late 1960s, when it was largely supplanted by the paradigm of Cognitive Psychology (Neisser, 1967) which returned to the direct study of humans. Despite this shift in the focus of the research, however, the common theme was the use of experimental research to understand the underlying mechanisms of human thought, learning and action, the development and refinement of theory.

This appeal to experimentation in the service of theory building is present today in education in general, and medical education in particular. One of the strongest recent proponents of the role of theory in education research has been Norman (1999). Norman and Schmidt (1999) for example, state that:

> ... as the history of natural sciences shows again and again, theory and understanding is a necessary precursor to progress, progress which leads ultimately to practical applications which were never envisioned by the discoverers. A theory ... is simply an expression of the relationship among variables, which shows, at its best, the precise quantitative contribution of particular variables in causing variations in others. It is hard for us to envision how one can achieve practical consequences of research without this kind of knowledge. *(p. 4)*

This general philosophy has driven much of the research on the nature of health professional expertise (Norman, Rosenthal, Brooks, & Muzzin, 1989; Papa, Stone, & Aldrich, 1996; Bordage & Zacks, 1984; McGaghie, McCrimmon, Boerger, & Ravitch, 1994), as just one example. These researchers have identified an interesting theoretical construct in another literature (such as cognitive, developmental or social psychology) and have attempted to generalize the theoretical predictions to the health professional education domain. In doing so, many have extended the original theory.

In hands of education researchers, however, the experimental tradition also developed a purely practical perspective regarding the capacity of the experimental paradigm to establish causality. That is, if we have control over a set of variables that cause an event to occur, then we can arrange conditions to ensure that the event will occur whenever we want. Thus, we can use our control of the causal variables to increase the frequency of events that we wish to occur and/or decrease the frequency of events that we do not wish to occur. This motivation is often the basis of applied research in education, which is directed at determining specific interventions that effectively improve learning. One of the most powerful commentaries on the use of this type of applied experimentation in educational research was provided by Campbell and Stanley (1966) who stated:

> This chapter is committed to the experiment: as the only means for settling disputes regarding educational practice, as the only way of verifying educational improvements, and as the only way of establishing a cumulative tradition in which improvements can be introduced without the danger of a faddish discard of old wisdom in favor of inferior novelties. *(p. 2)*

More recently, this strong position for the exclusive role of experimental research in the establishment of effective educational practice was reiterated by Colliver and Verhulst (1996) for the medical education field, who stated: "*In medical education, research is conducted to study links between teaching factors and learning outcomes.... Descriptive studies lack the powerful controls needed to establish credible links between exposures and outcomes*" (p. 211).

In fact, during a presentation at the American Educational Research Association, Colliver (1999a) argued not only that experimental research alone is suitable for use in performing applied education research, but also that applied research should be the only use of experimental research in education:

> ... we need to think more in terms of practical application, pragmatic consequences and outcomes. ... As part of this, we should reconsider the value of research that seeks primarily to understand ... [and] be wary of educational theory, such as cognitive theory, as a source of specialized knowledge about teaching and learning.

At its simplest level, this type of applied experimental research in education asks questions like, "Does this new innovation cause better learning?" The innovation itself may be fairly generic and transferable, such as the question of whether problem based learning is a better educational format in which to deliver content material (Colliver, 2000). Sometimes the innovation may be extremely specific, such as a question asking whether students benefit from interacting with a particular computer based learning package (Kaufman & Lee, 1993; Lechner, Lechner, & Thomas, 1999; Rowe, 1989).

At a more sophisticated level, applied experimental research may be directed at specifying the parameters of an educational innovation more carefully. For example, the research may try to find the limits of generalizability, the circumstances or populations for which the intervention is effective or not. Doucet, Purdy, Kaufman, and Langille (1998), for example, examined the use of problem based learning for the purposes of teaching active practitioners rather than students, asking in essence whether the use of problem based learning generalizes to continuing education contexts. Alternatively this type of applied research may try to identify the active ingredients in an already effective educational intervention, determining the components of an intervention that are the real cause of the better learning or that maximize learning. This was the type of applied research question that fueled the debate regarding the use of content experts as tutors in problem based learning tutorials (Dolmans, Wolfhagen, & Schmidt, 1996; Eagle, Harasym, & Mandin, 1992; Silver & Wilkerson, 1991).

The debate about the use of the experimental paradigm for theory building versus applied research is likely to continue. Several have argued that theory building has had a less vaunted history than has traditionally been promoted by the scientific community for philosophical reasons (Kuhn, 1962). On a more practical level, Plutchik (1968) has argued that attempts to test theories in psychology seldom produce clear-cut results because the theories themselves tend to be so vaguely defined that no definitive experiment can be performed. In his discussion, Plutchik cites Conant (1947), who suggested that a theory is never overthrown merely by contradictory facts, it is only overthrown by a better theory.

On an interesting note in medical education, it has been informative to follow the claims for the theoretical underpinnings of problem based learning as an educational innovation through the last 30 years. At its inception, the theoretical advantage of PBL was, at least in part, its ability to teach students to think like a doctor and develop general problem solving skills (Barrows & Tamblyn, 1980). As the psychological concept of general problem solving skills fell from grace and was replaced by concepts of context specific knowledge, however, the argument for the advantage of problem based learning shifted as well, with proponents now claiming that its advantage arises from the fact that the material is learned in the context of problems (Norman & Schmidt, 1992). The underpinning of the theoretical support appears to have shifted 180 degrees while the practice of PBL continues, largely unaffected by the debate.

It is also worth noting that the use of the experimental method for evaluating theory has often been plagued with a logical error, the error of confirmation bias in research. Custers, Regehr, and Norman (1996) have highlighted this logical error in their review of theories of knowledge representation in experts. Generally researchers in this field have some theory about the manner in which knowledge is represented in the expert's head (whether it is prototypes or instances or semantic axes or probability matrices). In testing their theory, they invoke the logical argument that, if their theory is correct, the expert should be able to perform some task better than a novice, or should be able to perform better on one task than on another. When the study is performed, the expected pattern is confirmed and the researcher concludes that this is evidence in support of the theory. However, this line of argument has actually committed the logical fallacy of confirming the antecedent. This type of logical error is not unique to psychology and education, and an historical example from epidemiology might help to clarify. Morrison (1960) in his discussion of the nature of the scientific understanding of causality describes a case of malaria. In medieval times, the prevailing theory was that malaria was caused by bad air (hence the name mal-aria) that tended to pool in the lowlands and swamps. If this theory were correct, then building villages on the tops of windy hills rather than in the lowlands should reduce the incidence of malaria. In fact, the expected pattern was confirmed in that there was a drop in the incidence of malaria in the hilltop villages, which was taken as evidence for support of the *bad air* theory. Today our theory suggests an alternate explanation for the benefit of

http://www.springer.com/978-1-4020-0466-7

International Handbook of Research in Medical Education
Norman, G.R.; van der Vleuten, C.P.M.; Newble, D.I. (Eds.)
2002, XIII, 1106 p. In 2 volumes, not available separately., Hardcover
ISBN: 978-1-4020-0466-7