Education Policy Analysis Archives

Volume 5 Number 13 **June 10, 1997** **ISSN 1068-2341**

A peer-reviewed scholarly electronic journal.

Editor: Gene V Glass Glass@ASU.EDU.

College of Education

Arizona State University,Tempe AZ 85287-2411

Qualitative Research Methods: An essay review

Les McLean
Margaret Myers
Carol Smillie
Dale Vaillancourt
Ontario Institute for Studies in Education/University of Toronto

Miller, Steven I. & Fredericks, Marcel (1994) *Qualitative Research Methods:*

Social Epistemology and Practical Inquiry.

New York: Peter Lang. 159 pages.

> …the more we try to unpack the notion of evidence for the sake of clarity, the more problematic it becomes, especially in trying to provide adequate justifications for educational policy making issues. (p. 119)

Abstract

The authors ask us to explore the topic of "qualitative confirmation" in relation to the processes and outcomes of qualitative research practice. The question that directs their inquiry is "how can we make a case that qualitative data or findings warrant the inferences about the topics we are studying?" We review the historical discussion of confirmation theory within the logic of

discovery, consider hypothesis generation and methodological decisions as instruments of the research process and then apply the Miller and Fredericks framework of rules to a published report of qualitative research (Glass, 1997). Full bibliographic references may be viewed by clicking on References (below) or on one of the linked citations in the text. We end our review with an appreciation of the work.

About the Authors

Les McLean

Les McLean is a Professor (Emeritus, as of July 1, 1996), Ontario Institute for Studies in Education of the University of Toronto (OISE/UT). He received his doctorate in Educational Psychology from the University of Wisconsin in 1964, specializing in statistics and research methods. After teaching at Columbia University Teacher's College, Dr. McLean joined the OISE faculty in 1966, teaching courses in measurement, statistics, quantitative and qualitative research methods and program evaluation. He continues to teach on a contract basis, year-to-year.

Research and development projects have included a national study of the evaluation of student achievement for the Canadian Education Association, several province-wide surveys of student achievement, direction of Ontario's participation in the Second International Mathematics Study and research into mathematics/language relationships in curriculum and pedagogy. Les, Doris Ryan and Barbara Burnaby directed an evaluation of four projects in the Canada/China Human Resources Development programme for the Canadian International Development Agency.

Publications include, *The Craft of Student Evaluation in Canada* (Toronto:CEA, 1985), *Learning About Teaching from Comparative Studies* (Toronto:Min. of Educ., 1987, with Richard Wolfe and Merlin Wahlstrom), and "The U.S. national assessments in reading: reading too much into the findings" (*Phi Delta Kappan*, 69, 5, 1988, 369-372, with Harvey Goldstein). "Time to replace the classroom test with authentic measurement" (*Alberta Journal of Educational Research*, 36(1), 1990, 78-84), "Student evaluation in the ungraded primary school: The SCRP principle" (*Proceedings of the Second Canadian Conference on Classroom Testing*, D. Bateson, Ed., UBC, 1992) and "Pedagogical relevance in large-scale assessment" (*Advances in Program Evaluation*, R. Stake, Ed., 1991).

For recent publications,see http://www.oise.utoronto.ca/~lmclean/>.

Les McLean, Professor-Emeritus lmclean@oise.utoronto.ca
Measurement and Evaluation (416) 923-6641, ext 2478
Department of Curriculum, Teaching and Learning
OISE/Univ. of Toronto
252 Bloor Street West
Toronto, Ontario, CANADA M5S 1V6

Carol Smillie

Associate Professor Dalhousie University School of Nursing. Vice President Canadian Cancer Society. Board of Directors of the National Cancer Institute of Canada and the Canadian Cancer Society. Director of the Sociobehavioural Research Satellite Centre of the National Cancer Institute's Sociobehaioural Research Network. Practice of Nursing has been concentrated in the area of Community Health and working with the Volunteer Health Sector.

Margaret Myers

Trained as Registered Nurse at St. Martha's School of Nursing, Antigonish, Nova Scotia; Bachelor of Science in Nursing from University of Western Ontario; Masters of Arts in Education from St. Francis Xavier, Antigonish, Nova Scotia; Have practiced in a variety of settings, including Nursing Education and Nursing Management. Speciality area is Program Developement and Implementation. Chairperson of the Diabetes Nursing Interest Group with the Registered nurses Association of Ontario. Am currently writing a textbook on diabetes education for health care professionals. Enrolled in Doctoral program in Curriculum at OISE at the University of Toronto. Married and living in London (Ont.)

Dale Vaillancourt

Dale's main interest is in using research to enhance online teaching and learning experiences, especially for women entrepreneurs. She is using her 25 years in the telecommunications industry and a full-time academic studies program in computer applications to help her meet this goal. For further details, see: http://www.oise.utoronto.ca/~dvaillancourt/resume.htm

---

The World Wide Web address for the *Education Policy Analysis Archives* is http://olam.ed.asu.edu/epaa

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-2411. (602-965-2692). The Book Review Editor is Walter E. Shepherd: shepherd@asu.edu . The Commentary Editor is Casey D. Cobb: casey@olam.ed.asu.edu .

EPAA Editorial Board

Purpose of our Review

The purpose of our review is to engage the reader in further discourse on this topic of qualitative confirmation. We will use Miller and Fredericks' (M&F) work as a vehicle to initiate debate and demonstrate its usefulness for critiquing published qualitative research studies.

Review Method

Following a review of the historical roots of confirmation theory and a description of M&F's method for establishing qualitative confirmation, we will apply the method to a journal article by Sandra Rubin Glass (1997), an extensive investigation of teacher and principal autonomy in both private and public schools. Her article is an attractive example because she offers readers ready access to much of her original data.

Historical Roots of Qualitative Confirmation

From a historical perspective, M&F take advantage of the *logic* provided by positivist theories of confirmation (e.g., Carnap, 1962, Hempel, 1965; Swinburne, 1973) and apply it to qualitative research. The authors argue that the use of a hypothesis and numerical probability statements continues to lend weight and credibility to conventional research and they seek to persuade us to transfer elements of the positivist paradigm to qualitative research. They recognize that "confirmation theory" involves complex issues and its application in the postpositivist arena is not a simple matter. Swinburne(1973) claimed that confirming one's findings must be reduced to probability (computational) statements, but M&F suggest that reducing qualitative research to quantitative calculations would overlook the advantages that naturalistic inquiry has to offer. From Carnap (1962), the term "confirmation" involves issues of classification (what is evidence), quantity (how much is good evidence) and comparison (what is the degree of firmness among the evidence).

In M&F's eyes, this raises important questions: whether numericallybased evidence is an option and whether the term confirmation can have different connotations. From Hempel's (1965) constitutive and regulative rules of confirmation, M&F recognize that hypotheses can be further supported by propositions if the propositions are defined legally and/or operationally. Again borrowing from the positivist school, issues related to indeterminacy and incommensurability (Quine, 1963, p.5,9) are introduced. If there is no objective reality and several data collection methods are allowed, triangulation is just as likely to yield contradictory views as it is to produce supportive ones. Which translation manual (2) should one use (indeterminacy)? Similarly, if competing translation theories inform the observations, and therefore cannot logically be used to determine which theory is correct (incommensurability), what does one use to make a decision? Charting the logic underlying one's inquiry process lends validity or correctness to the inferences (see also LeCompte & Preissle, 1993). They do not require translation manuals congruent with one's epistemological and ontological assumptions, but instead prompt the researcher to outline what other translations are possible and

determine which one or ones would be valid and why. Most qualitative researchers select their preferred method of inquiry, make a general statement about the data and show a few excerpts; this does not mean, however, that the selected process or the inferences are valid. M&F believe that in every study there is an *a priori* or *a posteriori* hypothesis, whether explicitly stated or nor, and that there can and should be a demonstration of how one's epistemological and ontological assumptions, triangulation approach, translation manuals and weighing of evidence link the evidence to that hypothesis. They adapt Hempel's rules to qualitative research (Hempel , 1965) and introduce the term, "qualitative confirmation."

Definition of Qualitative Confirmation

M&F define the term as: "those logical conditions that must obtain between the evidence and hypothesis" (p. 11). The authors admit that the term is an "oxymoronicsounding label" (p.1). However, they argue that their adaptation of the Hempel rules helps decide whether the logic underlying the methods chosen and the sampling and the weighing system for the evidence satisfies the conditions of qualitative confirmation. The rules are also expected to help one minimize the influence of personal bias. M&F recommend a systematic way of conducting and critiquing qualitative research that makes explicit the researcher's process from start to finish and highlights where pitfalls can be prevented. As already noted, one begins with an a priori or a posteriori hypothesis, which may be a transformation of the research question. This emphasis on a hypothesis would seem to set M&F apart from other qualitative research theorists (e.g., Glaser and Strauss, 1967; Stake, 1995), but we think not. See, for example, Strauss and Corbin (1990) on "The Research Question" (pp. 36-39), and Stake, Chapter 2 (pp. 15-33).

The overt thinking that guides the process and the resulting evidenceinstances (3) will ultimately be used to support or reject the hypothesis. This act of specifying one's logic, guided by rules, touches all phases of the qualitative research process; it is a demonstration why the planned activities and resulting evidence support the hypothesis. M&F state: " In other words, qualitative data have to 'demonstrate' their utility as potential evidential candidates for confirmation. It is not simply a matter of amassing positive evidence-instances for a hypothesis, but also showing *why* they contribute to supporting a given hypothesis" (p. 33, emphasis in original). Thus, qualitative confirmation provides a way of reporting the researcher's thought processes, helps promote a constructivist approach to qualitative research and answers the demand for increased rigor that Miles and Huberman (1984) identify: "Despite a growing interest in qualitative studies, we lack a body of clearly-defined methods for drawing valid meaning from qualitative data. We need methods that are practical, communicable, and not selfdeluding; scientific in the positivist's sense of the word, and aimed toward interpretive understanding in the best sense of that term" (p. 21).

Return to the Contents Table.

Go on to the next section.

Description of the Book

This is a scholarly book, in conception and content (but alas not in execution). We mean by "scholarly" that it digs deep for core concepts in what has been written about qualitative research methods-ideas, constructs, images that help us understand more than one case study, more than one ethnography, more than one narrative. Do such "core concepts" exist? Miller & Fredericks (M&F) think so. Otherwise how could they organize a book around the construct "qualitative confirmation"? What could it possibly mean to "confirm" findings from qualitative research? Reading Smith and Heshusius (1986) or Wolcott (1994), one might doubt the possibility.

Then there is postmodernism: doubt, doubt, doubt; everywhere we look there is doubt. A few years ago, the head of the British Museum of Natural History reviewed challenges to Darwinian theories of evolution and found merit in some of them. Attacked by critics for sowing doubt, he replied, "Doubt is splendid stuff." (4) Doubt may indeed be splendid stuff, but we humans seem to crave something else, something more solid in our mental life--something such as confirmation.

M&F offer material that is relevant to researchers everywhere but their editors have done them a disservice by presenting the work in a careless and awkward way. The lack of bold faced type and subheadings make an already difficult text even more difficult to read. References are given at the end of each chapter and each appendix, except for Chapter Three which has none. As a result, several are missing and there are many duplications. Outright errors appear that would be caught by even a cursory glance from an editor. (5) A curious feature of the book is that of its 153 pages, 62 are devoted to three appendixes which contain crucial elaboration or explanation. That said, we believe these authors provide a unique and powerful contribution to qualitative research--concepts and procedures that are worthy of study. *Confirmation* is the central concept of the book, a concept M&F argue, and we agree, has not been adequately addressed within the qualitative field.

Chapter One

Confirmation theory is introduced as the "major issue in qualitative inquiry", defining it as a description of what it means to say that the evidence relates to the hypothesis in the qualitative process. It is "the theory of when and how much different evidence renders different hypotheses probable" (Swinburne, 1973, p vi). It is acknowledged that this process is not free from assumptions and thus evidence can ultimately only be known through the senses. What constitutes data depends largely on who is looking for it; where mathematics may deal with pure fact, physics may not. The concept of weighing the evidence is introduced and *weighing* is distinguished from *weight*. Weight refers to the outcome while the weighing is the process of arriving at the outcome. In addition, weight can only be assessed by giving a rationale for the weighing of the evidence. Although M&F concede the concerns that have arisen over Hempel's classical framework of rules (1965), they value it and build directly upon it. Hempel defined: hypothesis, observation report,

observation sentence, entailment, direct confirmation, confirmation and development of a hypothesis for a class of individuals. (quoted by M&F on pages 8 and 9). They also cite Carnap (1962), who argued that although confirmation is central to understanding the evidence relationship, the term is also susceptible to different interpretations. Carnap, however, relied on computational interpretations and M&F argue that that would defeat the purpose of qualitative confirmation. The authors' intent is to use a non-probability framework to focus on a logic of discovery while not excluding the possibility of validation.

Since evidence comes in so many forms, rules are needed to help assess the evidence in relation to the research question (hypothesis). It is also necessary to develop a series of related but separate rules which bear only on the characterization of the data itself (p.12). M&F suggest that the ways qualitative researchers make a case for their findings will depend on two major aspects: the *labeling of* and the *justification for* using the rules. Two categories of rules are required: *constitutive* (defining what counts as a social situation or practice) and *regulative* (prohibiting or prescribing actions in situations defined by constitutive rules-taken from Greenwood, 1989).). In their discussion, M&F are "trying to understand what it means to be 'rational' in pursuing the activities of qualitative inquiry" (p. 14), and they reach back to "camps" defined by Wittgenstein, Popper and Donald Davidson (1984) for insight into rationality. To the present reviewers surprise, they then add "sociologists" to the list, a camp that "tries to determine the extent 'natives' hold their beliefs even in the face of (supposedly) other more 'rational' beliefs" (p. 14). Sociologists?

The last "camp" specializes the thorny and recurring issue of "translation", an issue that permeates any discussion of qualitative research. Whenever we use words to communicate the behaviour and/or feelings of others, the communication involves translation even when we all seem to be speaking the same language. Translation is discussed in every chapter of M&F, from many perspectives. Translation issues revolve around the ideas of indeterminacy and incommensurability. These two constructs are not dealt with at length. The "indeterminacy thesis" is: "because theories of meaning (i.e. those referring to natural languages) are not concerned with 'the fact of the matter', as are scientific theories in the natural sciences, it is possible to derive multiple translation manuals which may be incompatible but adequate for interpreting the behavior in question." The "incommensurability thesis" suggests that there is not even one translation manual that is adequate, as when trying to choose between two competing theories, for example (p.15). (For a comprehensive discussion of these concepts see Quine, 1960.) Translation is often at the level of data analysis rather than confirmation. Rationale for data collection processes and the subsequent description or translation of the data often is directed by the chosen research methodology. M&F close Chapter One with six statements describing how the above considerations relate to the notion of qualitative confirmation.

Chapter Two

The second chapter is devoted to the subject of hypotheses in qualitative research, not surprisingly since hypotheses are central to their entire approach. Hypotheses are viewed broadly "as statements which direct inquiry, in a Deweyan sense, in

relation to a theoretical framework, but without the necessity of such hypotheses being strictly deducible from such a framework" (p. 21). A priori hypotheses may emerge where the researcher has a theoretical framework in mind, for example, beginning with a category of behavior and linking it to a category of individuals. In contrast, the researcher may begin with an interest in a category of individuals and subsequently seek to link them to a category of behavior. M&F argue that both are qualitative because they are formulated in the context of a qualitative study and because their confirmation is dependent on qualitative data; they dodge the obvious tautology by acknowledging that both types can be formulated as well in quantitative studies. Confirmation, they say, is central to all research activities, but in qualitative research it is based on plausible logical relations which apply first to the data and then to the hypothesis. Qualitative research seeks to understand human behavior "from a perspective which, methodologically, requires qualitative data" (p. 25). Rules, they say, are needed if a case for qualitative confirmation is to be made. Since rules must apply both to data and hypothesis, M&F undertake to define what they mean by data-and their definition is entirely conventional: field notes, interview data, historical accounts and the like (pp. 26-7). The term "evidence-instance" is introduced to describe a discrete item of information to be put forward as part of a confirmation argument. A long elaboration on the meaning of "evidence" is presented in Appendix B, including more discussion how data become evidence. (6)

They discuss triangulation, breaking no new ground, and then tackle the really tough question: what counts as evidence and how do you judge its importance? Nothing new here, and we are all reassured to read, "It becomes quite difficult, then, to develop hard-and-fast rules for determining appropriate 'weighting' procedures for these kinds of issues" (p. 30). Some considerations that "may be helpful" include:

- The weight of the data-evidence is not necessarily synonymous with the "amount" of the data-evidence. The term "weight" can be loosely translated as an "absence of negative cases," that is, disconfirming instances. …
- Where the term "amount" of evidence is used, the implication is that some of the evidence must consist of (negative) disconfirming instances. …

M&F assert that contrary findings do not necessarily disconfirm the original hypothesis, for they may serve to clarify a previously unknown dimension. The matter of confirming vs. disconfirming instances is discussed more fully in Chapter Three. They tell us that using more than one data set can enhance the validity or reliability of data if it can be shown that the data sets are relevant to the problem--however, they neglect to explain how this can occur. M&F relate the purposes of using a triangulation process but the inherent problems associated with it are listed and not explained. We recommend that those unfamiliar with the process of triangulation consult other sources.

In what must have been an afterthought, the final paragraph of the final section, "Conclusions", introduces a muddled discussion of the terms "relationship" and "association". Hopes that the muddle will be cleared up are not fulfilled.

Chapter Three.

In chapter three M&F present their four rules for qualitative confirmation and instruct us in the use of these rules, including how to handle negative cases or cases of disconfirmation. Overall, it is the most carefully constructed and clearly presented chapter. How qualitative findings may become evidence for a hypothesis or research question is the focus of the discussion here and the authors impress us with how their qualitative research rules apply to both sides of the research continuum. One advantage claimed for researchers is that the rules allow for the development of qualitative research studies with confirmation in mind. They also suggest how research findings are to be interpreted when choosing data sets. As for qualitative disconfirmation, some additional advice is given in the interpretation of negative evidenceinstances. Researchers usually have to use the qualifier "some" instead of "all" (QuineDuhem Thesis) because scientific theories may be composed of auxiliary hypotheses for certain predictions (p.48). Auxiliary hypotheses can allow incompatible evidence while still not disconfirming the original hypothesis. The null hypothesis also finds a place in Miller and Frederick's qualitative confirmation theory.

Chapter Four.

Chapter Four moves into the practical application of rules to five articles published since 1980. M&F pick up on a suggestion from Miles and Huberman (1984) that researchers ask systematically a series of questions before, during and after the research: (p.53): . what is (are) the major research question(s) or hypotheses? . what method(s) will be used and what sort of data yielded? will other(indirect) data sources be utilized? . how will the case be made that the research question/hypothesis has been or has not been confirmed? . if mixed data sources are used, how will they be handled in terms of confirmation/ disconfirmation? All researchers are advised to include a final section in the research report (possibly as an appendix) which explicitly addresses the issue of qualitative confirmation. Such a section helps the researcher focus on the larger epistemologicaltheoretical questions raised by the study and reminds us that the presentation of our findings can be regarded as a type of translation manual as well as a demonstration of the rigor in one's research.

The application of confirmation theory to the five caseshelps a little to understand M&F's method, but their choice of articles left much to be explained. A graduate student listed 20 reports of qualitative research from the years 1970-1990. The same student selected 7-10 for further study, in an attempt to present a "rough cross section" of work. In our view, a more qualitative approach would have been preferable: purposive choice to illustrate application of the rules rather than this version of sampling. In only one study, for example, was there any attempt to discuss negative or disconfirming instances. In most there was no attention to the issue of confirmation, but in one or two there was an implicit suggestion that confirmation had been achieved. There was no discussion of the weighing the evidence in making qualitative claims. Given the rather radical nature of M&F's proposals, it should perhaps come as no surprise that research conducted and published more than a decade ago does not adhere closely to the Rules. In order to help future researchers, M&F provide a "Checklist for Qualitative Confirmation" at the end of the chapter, including the questions mentioned above and ending with "What, after all, has one discovered/concluded from this investigation? Is it truly

warranted?" (p. 66-7)

Chapter Five.

The final chapter (before those three long appendixes, that is) is entitled, "Epistemological Asides and Conclusions"; fearlessly confronting the big question of "truth". Suppose a study has met all the criteria for qualitative confirmation, "Is this all that is necessary for declaring that the findings, then, are 'true'?" (p. 73) M&F advise us to look at various philosophical camps, bringing Winch again (1958, 1964) and again invoking the "camps" (from Chapter One). Qualitative confirmation provides a framework for establishing qualitative data as evidence for the hypothesis, that is, it meets the requirement of rationality, but it is less successful as regards indeterminacy. It is not possible to rule out different or competing translation manuals.

As readers will have gathered by now, this concept of "indeterminacy" pervades the book--as it does the field of qualitative research. It is a way the philosophers have discussed whether there can be any "truth" and the way M&F discuss the prevailing view in qualitative circles that there is no truth-there are no methods that allow us to reduce indeterminacy to zero. Readers may feel, with justification, that M&F's discussion of the concept is fragmented and incomplete. Fear not; Appendix A (24 dense pages) is devoted to it.

Appendix A. Some notes on the nature of methodological indeterminacy.

Here the authors return to the intellectual roots of qualitative confirmation and elaborate the construct, "methodological indeterminacy". M&F revisit Quine and lean heavily on Roth's 1987 book, Meaning and method in the social sciences: A case for methodological pluralism. Roth, echoing Winch (1958), opens his book by referring to "the general collapse of positivism" and concludes (following Quine) that multiple translation manuals are not only possible but inevitable. There is no "fact of the matter" in relation to the human sciences.

Where can we go from here? Undaunted, M&F press on, seeking to dodge the philosophical bullet by concentrating on practical concerns, namely methods.

> …the genuine problem of indeterminacy for the human sciences does not lie at the level of debates concerning "hermeneutic" vs "scientific empirical" *views* of human action, but rather at the level of *specific* methodological techniques that ultimately are the constitutive elements for determining the existence (or lack of it) of indeterminant translations. (p. 93, emphases in original)
>
> ……
>
> To be clear, we are not arguing that the definition of methodology as the application of specific techniques and procedures is sufficient to resolve indeterminacy in all situations, but only that it is a necessary condition both to establish its existence and to demonstrate *possible* ways of reducing it. (p. 94, emphasis in original)

Some clarity does emerge: specific techniques and procedures (methods) are needed to *establish the existence of* indeterminacy. It is always there, we suppose, but until you apply the techniques and procedures (collect the data?), you don't see it Now that positivism has collapsed, we go farther: application of methods "may *generate* a type of indeterminacy" (p 101), or "types of indeterminacy *that are produced by* competing methodologies" (p. 103, emphasis added). This is surely correct, because different methods bring to light different amounts and kinds of indeterminacy-hence the term, "methodological indeterminacy". M&F borrow from quantitative methods and discuss differences "between" and "within" methodological approaches, with regard to both qualitative and quantitative approaches. The discussion of qualitative approaches is much more insightful than that of the quantitative, the latter naïve and verging on the trivial. We question M&F's assertion, citing Fuller (1988), that "there is a heavier 'burden of proof' on the qualitative side of the equation for showing either indeterminacy or lack of it." (p. 105) A better statement, in our opinion, is that for practicing researchers, "indeterminacy is usually perceived itself as a 'working hypothesis'; one which may never in principle be unequivocally 'accepted' or 'rejected' but one, nevertheless, that is capable of empirical inquiry." (p. 105)

Appendix A concludes with a short discussion of the status of theory in the human sciences. (M&F leave no topic completely untouched!) They cite a belief among "some quantitative researchers" that improvements in measurement will solve our problems, including give us better theories. In our opinion, they have it backwards: better theories lead to better measures (and the process recycles). Let M&F have the last word, with which we completely agree:

On the other hand, for those who do not see the possibility (or usefulness) of equating the human sciences with the natural sciences, the indeterminacy reflected by methodological applications can be "reduced" over time, but its reduction is directed towards making human behavior more "intelligible" rather than more "scientific". (p. 108)

Appendix B. Clarifying the "adequate evidence condition" in educational theory and research

The authors set out to convince us how important it is to clarify what they call the "evidence condition". Lack of attention to what is meant by evidence "has resulted in a restricted view of this term with a correspondingly unwarranted optimism regarding the formulation, implementation and evaluation of educational policies and practices." (p. 117) Faithful to their book title, they promise a thick description type of analysis to attain their goal. If "dense" is the same as "thick", then they deliver on this promise.

Israel Scheffler (1965) is cited early, as we would expect, and not just because he introduced the phrase, "adequate evidence condition". No discussion of evidence in educational research would be complete without at least a mention of this classic. New sources to us are Lakoff (1987) *Women, fire, and dangerous things: What categories reveal about the mind*, and Lakoff & Johnson (1980) *Metaphors we live by*, insightful discussions of the nature and process of category creation. Since

qualitative researchers live and die by their categories, these are helpful references for anyone (such as the authors of this review) who had overlooked them. The process of categorization is the formation of "Idealized Cognitive Models …culturally unique and semantically-based modes of perception and reasoning whereby users of a language construct cognitively-based models to explain social and physical reality." (p. 124) There are five types of models: Cluster, Metonymic, Social Stereotype, Ideal and Metaphoric.

We are not attempting a summary of this appendix. M&F describe the concept of evidence as having "entrenched ambiguity" and "overall complexity", and after reading the appendix several times we heartily agree. Lawyers and judges struggle with it daily, of course, and law schools have whole courses on it. M&F also write, "Paradoxically, however, the more we try to unpack the notion of evidence for the sake of clarity, the more problematic it becomes, especially in trying to provide adequate justifications for educational policy making issues." (p. 119) Here are a few points to give the flavour of the work.

Among the ambiguities cited is the "conflation" of the terms evidence and data, as we noted in the description of Chapter Two and in Note (6). According to M&F, data are *candidates* for evidence but only become evidence "to the extent that they are consistent with both the larger class of rules identified with the particular methodological approach, and with the rules regulating the application of a particular technique or tool within the methodological approach." (p. 118) Whew!

Educational implications are offered, including ambiguities surrounding the concept of "school effectiveness" and the lack of consensus on adequate evidence for effectiveness. You might not agree with everything they write, but you will be led to think hard about it.

Appendix C. Reciprocal paradigm shifts and educational research: A further view of the quantitative-qualitative dilemma

In the concluding appendix, M&F state their belief and concern that while those involved in educational research perceive a paradigm shift from quantitative to qualitative inquiry, many are not convinced that the qualitative-interpretationist view can be sustained on the basis of conventional scientific criteria. Some, Miles and Huberman (1985) for example, manage to sidestep this issue and some recognize qualitative research but believe that the eventual testing or proof of causality or theoretical prediction must be left to scientific empiricism (Popper, 1969, Rosenberg,1988). M&F concentrate their discussion on the interpretationist framework and they examine some commonly held assumptions that underlie the two research processes. They present their position that quantitative and qualitative approaches need not be viewed as contradictory in terms of some educational problems within some contexts, and a methodological mix is not only possible but desirable.

Miller and Fredericks do not agree with the asumption that scientific empirical data is quantitative data and that the constructs of validity and reliability are crucial issues only in quantitative studies.

One of our central claims will be that not only are the crucial issues of reliability

and validity perfectly general across all research based forms of inquiry; but also that in meeting the requirements of reliability and validity, interpretationist accounts may yield warranted conclusions that, while different in form and content from scientific empirical claims, can, nevertheless, be compatible with them. (p. 137)

They remind us that these assumptions have their roots in the traditional confirmation theory of Carnap (1962) and Hempel (1958) and are drawn from a "probabalistic-operational" perspective. This approach to the definition of scientific evidence seems to rule out the possibility of interpretationist accounts of data that can only be understood rationally as qualitative data. In response, they put forward and discuss the following assumptions (p. 139):

1. interpretationist accounts within the human sciences are correctly based on the direct or presumptive use of non-quantitative research strategies,
2. these strategies must be viewed as being fundamentally different in their qpplication and in the interpretation of the data they produce from those research strategies based on probability assumptions,
3. they are, nonetheless, subject to the general scientific constraints of reliability and intersubjectivity, and
4. they can be employed as alternative means for "confirming" hypotheses and/or providing for critical tests for theoretical perspectives.

These assumptions lead them to confident statements about the tough issues of causality and theoretical interpretation of reports from qualitative studies. "First, the qualitative research agenda stipulates only that causal connections (as either necessary or sufficient conditions) of behavior be meaningful in terms of the native's own accounts and that such connections be accurately described. Secondly, the investigator's theoretical interpretation of these accounts *can* incorporate 'higher order' or more general constructs, but without the implication that their use must *necessarily* 'reduce' the original accounts." (p. 145, emphasis in original)

Return to the Contents Table.

Go on to the next section.

Other views relevant to qualitative confirmation

This section is presented to place M&F within the larger qualitative research community. Our discussion is in no way a complete review of the voluminous literature; it highlights different interpretations of the hypothesis setting process, validty and reliability concerns and understandings of translation and triangulation methods. In this section, we emphasize with italics the words that seem to us to be versions of *confirmation*. In many ways, LeCompte and Preissle (1993) support M&F's concept of qualitative confirmation. LeCompte and Preissle agree, at least metaphorically, that validity is involved in many aspects of the inquiry process: "How validity is defined and treated varies according to what researchers do, what tasks they are undertaking, and in what phase or stage of the research they are in" (p. 325). Theoretical frameworks, general design, context, participants, researcher experience and procedures of data collection and analysis have a bearing on the issue of validity. As LeCompte and Preissle say, "Consequently, although we urge scholars to discover and formulate what their research philosophy is, we believe that it is only one factor contributing to how validity is defined" (p. 326). They also caution, as do M&F, that replacing qualitative processes with strictly quantitative ones erroneously prompts a single consolidated definition of validity and potentially jeopardizes richness of detail and creativity. For LeCompte and Preissle, qualitative research is idiosyncratic and data analysis entails an emergent process: "Even midway through an analysis,uncertainty and frustration accompany the unfolding direction" (p. 330). They see qualitative research as loosely connecting researchers who come from a broad spectrum of philosophical traditions; there is not just one. They could ask that all the different qualitative researchers state how their philosophies decide validity and then apply those guidelines to the study, but they stop short of this because they see it as an *a priori* assignment approaching "determinism" (p. 326).

A comparison of their descriptions of research phases demonstrates how they propose to obtain validity (or qualitative confirmation) especially in the areas of: formulating goals, developing a research design, selecting data sources, experiencing and directing the research, collecting data, collaboration, comparing phenomenon and data analysis. Part of developing research questions is to ensure that the research goals, the context of the situation and the interests of the stakeholders are aligned. M&F would say that this process entails what must also be considered when developing a hypothesis. Much of the discussion around validity stems from concerns about the sources that are assumed to provide validity. LeCompte and Preissle argue (as do M&F) that qualitative research methodologies cannot discount the range of ontological and epistemological assumptions and theories that are at their disposal. The key to validity is that researchers must be aware of and acknowledge the use of disparate views. The kinds of evidence colleagues accept as legitimate and adequate may be affected by who is being studied. More data may be required because of who is going to scrutinize the results. Therefore, who or what is to be studied must be considered when deciding on a design. This is similar to M&F's concept of weighing (but not 'weighting').

Will there be other extraneous factors such as background evidence that needs to weigh into the confirmation of the hypothesis? Researcher's background and role in

the investigation are central to how the validity is addressed:

> History teaches that attention to the individual researcher is relevant to validity in qualitative research. What background and training does the researcher bring to the investigation? How carefully, thoroughly, openly, and honestly are researchers known to do their work? Who was responsible for the researcher's training? What reputation has the scholar earned in previous investigations? What does the researcher report about participation in the research? Introspective and reflective amounts of influence on what is seen and heard contribute to the audience's confidence that the researcher attempted to track these factors." (p.329).

This is similar to M&F' constitutive and regulative rule that revealing the researcher's training and practice as well as the ethical or operational practices will help us assess how much *credibility* can be given to the researcher's work. LeCompte and Preissle agree with M&F that a systematic way of collecting data can be used to give more credence than one that is not. Both pairs also agree that just because something is done correctly, it may strengthen the research but does not necessarily mean that the results are sufficient to meet the criteria. LeCompte and Preissle suggest many ways for researchers to enhance confidence in their results, for example, through collaborative participation with the participant, congruency between theory and observation, intermethod and interobserver checks, personal reflection and introspection. Therefore, although LeCompte and Preissle would determine their analysis procedures at a different stage than M&F, they do agree with M&F that several options are available. Both agree that one should use multiple methods to reduce the possibility that bias will affect the credibility or validity of the results.

LeCompte and Preissle do not strive to provide conventional external validity because small sample sizes usually make this task impossible. However, they state that comparing phenomenon is useful and can be achieved by defining the "typicality" (Wolcott, 1973) of a phenomenon. Threats to comparing phenomenon are whatever obstructs or reduces a study's *translatability*. Translatability is the degree to which the researcher can adopt theoretical frames, definitions and research techniques accessible to and understood by other researchers in the same or related disciplines. Thus LeCompte and Preissle agree with M&F that alternative translations should be considered. However, LeCompte and Preissle think of translatability in terms of its usefulness in linking with others, whereas M&F think of translatability for reducing bias and confirming evidence or a hypothesis.

The one other area in which LeCompte and Preissle are distinct from M&F is in data analysis. LeCompte and Preissle state simply that one cannot predict what will happen, that trying to develop and use a template to direct the data analysis is impossible. They feel that "qualitative analysis is interpretive, idiosyncratic and so context dependent as to be infinitely variable. A creative analyst can never be sure that the ending will match the point of view adopted in the beginning" (p.330). In closing, LeCompte and Preissle agree with M&F that a single definition of validity is inappropriate for qualitative research. In a qualitative confirmation process, all authors agree that concerns about validity touch every part of the inquiry.

LeCompte and Preissle use triangulation to understand a phenomenon; M&F use triangulation to confirm a hypothesis.

Guba (1981) outlined several paradigms for discovering "truth". These include a judicial paradigm that has well established rules for procedure, rules of evidence and criteria for judging the adequacy of the rationale for a proceeding. This judicial paradigm offers guidelines for behaviour. Another paradigm is that of expert judgement. The third is what he refers to as the rationalistic paradigm and is essentially connected to deductive thinking and a logical positivist point of view. The paradigm he obviously prefers is the naturalistic. The naturalistic is characterized by inductive thinking, and phenomenological views of knowing and understanding social and organizational phenomena. He notes that there are shades of grey in viewing these paradigms and that often they are seen as competing but in the task of knowledge production they are all important. Guba stresses that the naturalistic ecological hypothesis is imbedded in a context which is often more powerful in shaping behaviour than differences among individuals. In conclusion, Guba states that understanding the reality of the world requires acceptance of the notion that the parts cannot be separated. He further concludes that because of the assumptions underlying naturalistic enquiry the traditional concerns for objectivity, validity and reliability have little relevance for the design of the research. The validity of the findings is related to the careful recording and continual verification of the data that the researcher undertakes during the investigative practice. This is consistent with Wolcott (1990,1994).

The questions of translatability and comparability trouble qualitative researchers (Goetz & LeCompte,1984,1993; Wolcott, 1973). Appropriate sampling improves the generalizability of quantitative studies but researchers improve the quality of their qualitative studies by (among other things) ensuring that units of analysis, concepts generated, population characteristics and settings are described and defined so that they can be compared between studies. Translatability is related to the degree to which the researcher uses theoretical frames, definitions and research techniques that are accessible to or understood by other researchers in similar or related disciplines (Goetz & LeCompte,1984, p228). M&F do not hold contrary views to these discussions of scientific validation; rather they have built upon them in suggesting the possibility of confirmation of findings. Lincoln and Guba (1985), in outlining their fourteen characteristics of operational naturalistic inquiry, write: "naturalistic ontology suggests that realities are wholes and cannot be understood in isolation from their contexts, nor can they be fragmented for separate study of the parts; because of the belief in complex mutual shaping rather than linear causation, which suggests that the phenomenon must be studied in its full scale influence, and because contextual value structures are at least partly determinates of what will be found" (p.39). Spindler and Spindler (1992), when developing their eleven criteria for good ethnography say in Criterion II, "hypotheses emerge in situ as the study continues. Judgments on what may be significant to the study is deferred until the orienting phase of the field study is completed." M&F dodge these issues for most of their book, but they confront the issues in Appendix C.

On the subject of verity and what constitutes rigor, Marshall and Rossman (1989), suggest the use of "controls" which rely heavily on other researchers: the use of a research partner to play devil's advocate; a constant search for negative instances

(from Glaser and Strauss, 1967); checking and rechecking the data, and purposeful testing for rival hypotheses; practicing "value free" note taking, and using "strictly objective" observations; devising and applying tests to the data; using the guidance of previous researchers to control for data quality; and conducting an audit of the data collection and analysis strategies (Lincoln and Guba, 1985, p.148). Marshall and Rossman neglect to tell us how a research partner will play "devil's advocate", explain what constitutes "value free" and "strictly objective" observations--nor do they tell us what types of tests to develop, how to develop them, or at what points we should test the data. Negative instances feature in M&F's rules, but their discussion is thin compared to others such as the above.

Talbot (1995), lists a number of factors involved in the *credibility* of findings: remaining in the field over a long period of time; using triangulation; negative case analysis; and having participants review researcher's interpretations and conclusions. She would appear to be in M&F's camp, because she uses the term "confirmability" to describe the process where findings, conclusions and recommendations are supported by the data, and she suggests that there should be an internal agreement between the investigator's interpretations and the actual evidence. Confirmability is only one of four factors that establish *trustworthiness* for Talbot, however, the other factors being credibility, transferability, and dependability. In establishing credibility, she borrows from Goetz and LeCompte (1984) who say that "establishing validity requires determining the extent to which conclusions effectively represent empirical reality, and assessing whether constructs devised by researchers represent or measure the categories of human experience that occur." This definition leans far too heavily on quantitative means in calling for a representation of empirical reality but is explicitly written and comprehensive in description in all other respects. *Credibility* crops up again as Lincoln and Guba (1985), discuss four constructs against which the *trustworthiness* of a study can be evaluated: credibility; transferability; *dependability* and (here's that confirmation again) confirmability. Credibility depends on how accurately the subject is identified and described. Transferability is noted to be impossible from the stance of external validity, but is greatly assisted by providing the greatest possible range of information, and thick descriptive data. The applicability of one set of findings to another setting rests more with the later researcher making the transfer than the original researcher. The authors point out that dependability is difficult to predict in a changing social world. In establishing dependability, the researcher attempts to account for changing conditions in the phenomenon chosen for study as well as changes in the design created by increasingly refined understanding of the setting.

The preface to Strauss & Corbin's (1990) book puts the question of confirmation clearly up front for those who are interested in inductively building theory through qualitative analysis of data. They state that however exciting may be the experience of gathering data, there comes a time when the data must be analyzed and at that time the researcher asks, "How can I have a theoretical interpretation while still grounding it in the empirical reality reflected by my material?" "How can I make sure that my data and interpretations are valid and reliable?" and "How do I pull all of my analysis together to create a concise theoretical formulation of the area under study?" The research question in grounded theory is a statement that identifies the phenomenon to be studied. As the researcher proceeds through the process they can

go in different directions and therefore the questions can change. M&F seem able to accommodate this definition of research question and would assure the grounded theory researcher that what needs confirmation is the final conceptualization of the proposed theory and whether or not the symbols of behaviour and language to which the researcher reacted were true. The conceptualization question is one of confirmation; the latter is a question of methodology. Anselm & Strauss strive for *rigor* with seven evaluation criteria embodied in the questions: (1) How was the original sample selected? (2) What major categories emerged? (3) What were the indicators that led to the development of the categories? (4) What categories directed the theoretical sampling process? (5) What were the hypotheses pertaining to conceptual relations among categories and how were these tested? (6) Were there instances in which the hypothesis did not hold up to what was seen (7) How and why was the core category selected and on what grounds were the final analytic decisions made?

We can see within these criteria many of the issues of confirmation and disconfirmation addressed by M&F. They also relate to the criteria for the evaluation of qualitative research proposed by Lincoln and Guba(1985), who were among the first to suggest criteria for good qualitative research. Strauss and Corbin(1990) and Lincoln and Guba (1985) differ in their area of emphasis, with Strauss and Corbin(1990) being more concerned with internal validity and Lincoln and Guba(1985) wanting the work to shed light on other instances. M&F seek to accommodate these different dimensions and to add a strategy that will enhance work that has been started by others.

Return to Contents Table.

Go on to the next section.

Rules for Confirmation

Qualitative confirmation rests on the assumption that there is a hypothesis or at least a hypothetical statement for the study, that data collected through qualitative methodologies will remain qualitative data and that this data can be manipulated within a set of rules. M&F propose just such a set of rules for confirmation of qualitative research. The rules are the application of deductive reasoning in an effort to confirm that the research data logically confirm or disconfirm the hypothesis framed for the study. To apply the rules, therefore, the researcher's first task is to define a hypothesis; it can be either *a priori* or *a posteriori*. The next challenge for the researcher is to determine data collection methods that will produce "evidenceinstances". "Evidenceinstances" are the result of the qualitative data being recast as evidential statements for confirmation. Here are the rules, exactly as written by M&F (pages 4142, emphasis in original):

- **Rule 1**: The qualitative evidence instances must be positive for the development of the hypothesis. If they constitute a *denial* of the hypothesis, they disconfirm the hypothesis (see Hempel's 9.3 Df.).
    - 1.1 The limiting "weak" case for confirmation would be the existence of only one positive instance, and for disconfirmation the existence of one (and only one) negative instance.
- **Rule 2**: If the evidence instances constitute a methodologically unique class (8), they must (minimally) not be contradictory to one another.
    - 2.1 As a class, the statements should entail the development of the hypothesis, while the possibility of their own entailment(s) is left open.
    - 2.2 If this class consists of only two instances and they contradict each other, then the hypothesis is neither confirmed or disconfirmed.
    - 2.3 If this class consists of numerous instances, *some* of which are contradictory to the hypothesis, then (2.2) obtains, unless it can be shown that there are *more* instances (positive or negative) and these should be counted for or against the hypothesis, or the above instances should be given a priori "weights" in terms of importance. Note: the assignment of these weights could be given by an agreement of knowledgeable experts.
- **Rule 3**: If relevant background evidence can be adduced for the hypothesis under consideration, and if this evidence alone is sufficient for the development (e.g., entailment) of the hypothesis, and, furthermore, if it is non-contradictory to the class chosen as evidence for confirming the hypothesis, then the given hypothesis may be said to be confirmed.
    - 3.1 If the background evidence is sufficient for the development of the hypothesis but is contradictory to the evidence chosen for confirmation, the hypothesis' confirmation will remain undetermined.
- **Rule 4**:For a given hypothesis that is to be confirmed by evidence instances derived from a variety of methodological approaches, the class comprising these statements should first be partitioned into relevant categories, i.e., historical-narrative, ethnographic, documentary, quantitative, etc.
    - 4.1 statements within categories should be internally consistent (non-contradictory)

- 4.2 Each partitioned subset need not be sufficient for the derivation of the hypothesis, but the totality of subsets comprising the class should be sufficient (and, hopefully necessary).
- 4.3 Depending on the number of subsets, if one subset contradicts the others, either partially or wholly, the confirmation of the hypothesis is left undetermined.
- 4.4 If one subset is disconfirming to the hypothesis but neutral or non-contradictory tothe remaining subsets, and the remaining ones are consistent, the hypothesis will be considered confirmed.
- 4.5 The hypothesis will be considered disconfirmed (i.e., rejected) if (a) all subsets disconfirm, (b) if a majority disconfirm, or (c) in the limiting case of two subsets, where one is contradictory to the other, if additional grounds (i.e., agreement by experts) can be adduced for the disconfirming subset, this will be taken as sufficient for disconfirmation. (In this case the evidence has been "weighted" towards disconfirmation; of course, the converse, i.e. for confirmation, could be similarly argued.)

Return to Contents Table.

Go on to the next section (Application of Rules).

Application of the M&F framework to the Sandra Rubin Glass study, "Markets and Myths: Autonomy in Public and Private Schools", published in *Education Policy Analysis Archives*, vol. 5, no. 1, January 6,1997 (the 'Glass' study)

The title describes the article well: Glass undertook a large-scale study of teacher and principal autonomy in public and private secondary schools and used the data to explore claims made by Chubb and Moe (1990). She challenges the latter's assertions: that the organization of private schools offered greater teacher autonomy resulting in higher student achievement and that the bureaucracy of public schools stifles autonomy and limits student achievement. Glass attempted to bring to the surface conditions which constrain teacher and principal autonomy in both public and private schools. Although she did not express it this way, it is reasonable to state that she had a hypothesis: "There are no differences between public and private schools in the amount of autonomy teachers and principals have." She also set out to answer the question, "What conditions impede teacher and principal autonomy in both public and private schools?", implying the hypothesis: "There are no differences between public and private schools in the conditions that foster or inhibit autonomy."

In her comprehensive study, Glass used both data source and methodological triangulation. "The methods employed in this investigation were those of the multi-site qualitative case study: interviews from multiple data sources, observations and field notes from a variety of on-site meetings and visits, and analysis of documents (brochures, teacher handbooks, policy manuals, meeting agendas)." She conducted an intensive study of three public and three private secondary schools, interviewing fourteen private school teachers, fifteen public school teachers; assorted principals, heads and assistants from each school were interviewed at their respective sites.

According to M&F, triangulation is described as a series of strategies that directs both the generation of data and the clarification of findings . They go on to say that the purpose of triangulation in qualitative inquiry is "to provide a rationale for increasing the plausibility of qualitative findings" (M&F, p.28). M&F state that, "if different data sources are used for the study of a particular problem, and if it can be claimed that they are relevant for the problem, the likelihood for the total data set to reflect reliability and validity will be enhanced" (p.27). The different sources listed by Glass have face relevance to the problem and in principle enhance the likelihood of reliability and valididy. In the journal article, however, the only data we can identify by source is from the interviews. A set of hyperlinks allows readers access to them as the findings are being reported. Whenever a quotation is given from an interview, the reader may choose to examine the quoted passage in its full context by clicking with the mouse on an icon at the margin. (Glass has taken the extra step of directing the link to the exact location of each quotation in the interview and highlighting the quotation in bold.) We thus have explicit examples of the "evidence-instances" referred to in M&F's rules and an unprecedented opportunity to assess in detail whether the evidence is "positive for the development of the hypothesis". The thirtyseven interviews are presented in a standardized pattern, clear and accommodating to read. She used many of the

interview questions from the Moles (1988) and Blase's (1991) surveys (part of the 'High School and Beyond' survey) used by Chubb and Moe (1991) to develop their index of teacher and administrator autonomy. The use of structured tools that have been previously tested lends credibility to the data set. Participants were selected based on years of teaching experience (at least five), and years of experience at the present school (at least three). The school sites were chosen so that the constituencies of both the public and private schools were as comparable as possible; in both private and public situations, schools servicing high income families and focusing on academic excellence and college preparation were selected as well as less favoured districts. Glass does not explicitly state how the data will be used in the confirmation process (she did the work before M&F's book was published), but she does share her process of weighing the evidence. For M&F, this consideration is a necessary step in research design that allows for consequent evaluation in relation to qualitative confirmation.

We now assess the Glass study for qualitative confirmation by using the M&F rules.

As noted in the previous section, **Rule 1** states that the evidence instances must be positive for the development of the hypothesis. If they constitute a denial of the hypothesis, they disconfirm the hypothesis. In more than thirty interviews with teachers and their principals, from both private and public schools, Glass found that participants from both private and public schools experienced about the same measure of autonomy in their environments or were able to work around conditions that constrained it. The interviews brought out the degree of complexity inherent in the idea of autonomy. Rule 1 is thus satisfied.

Rule 2 states that if the evidence instances constitute a methodologically unique class, they must (minimally) not be contradictory to one another. The interviews constitute just such a class, and the evidence instances Glass presents do not contradict each other. In most journal articles, or even books, we have to be content with the evidence instances provided by the author-always and necessarily a small subset of all possible instances-but here we are able to read all the text of all the interviews. We do not claim to have done so, but each one of us read at least one interview through, explicitly searching for evidence instances that would contradict the hypotheses. We found none. While the researcher does not report explicitly on the data realized from onsite meetings, visits, and analysis of documents, she does refer to the high achievement standards of schools, curriculum content, parent involvement, and how arrangements were made to collect data--information that had to be collected from these sources. Consider, for example, the following statements from Glass's "Findings" (we have put inferences assumed to be made from field notes and the like in bold):

In a private school, new teachers will generally define the curriculum predicated on their own content knowledge and interest. Because of smaller faculty numbers, there may be two or three other teachers with whom to coordinate curriculum; yet each teacher specializes in a particular facet of that content area. While each of the three independent schools in this study have either a middle school or middle and elementary school as part of its organization, **students come from a variety of other schools**. Consequently, **coordination is a**

matter of interest only within the upper school. Any coordination of curriculum is accomplished within the institution, as described by this private school teacher: (a quotation follows, with link to the interview).

……

This study was conducted in a right-to-work state in which teacher unions are virtually non-existent, **but teacher associations are predominant**. These **associations are seen as variously strong or weak depending on locale**. Only **one of the three public schools is in a district having a very strong teacher association. Most, if not all, of its teachers are members of the association and quite a few are active in its leadership. The other two schools are in districts that negotiate teachers' contracts with the association, although the faculty are much less active.**

It is not possible to confirm or disconfirm these statements from the journal article, because the relevant evidence instances are not supplied. Those that are supplied satisfy Rule 2, but what do we say about the others? We do not know.

The Chubb and Moe (1990) findings are presented, and we could regard them as "background evidence" (**Rule 3**), but rather than being "adduced for the hypothesis", they are findings to be questioned, and possibly to be disconfirmed. This is an outcome of research that M&F appear not to have foreseen. Glass's opening statement, and her analysis of data is an argument for a disconfirmation of their report. On the other hand, background information from Sedlak (1986) and Ball (1987) are adduced as support for Glass's hypothesis. In our opinion, the evidence of Sedlak and Ball is not sufficient for development of the hypothesis, but neither are they contradictory. Applying Rule 3, we would say the hypothesis is confirmed by Sedlack and Ball but contradicted by Chubb and Moe. The hypothesis would therefore remain as "undetermined" if the Chubb and Moe data were sufficient. This is why Glass argues so strongly that the Chubb and Moe evidence is, to put it mildly, not sufficient. She points out that they present no evidence whatsoever on private schools from the "High School and Beyond" study that is the basis for their arguments. Applying Rule 3, we conclude that the background evidence is itself contradictory and would leave the hypothesis undetermined. It seems as if we have to decide for ourselves how to weight the evidence.

Rule 4 states that, "for a given hypothesis to be confirmed by evidence instances derived from a variety of methodological approaches, the class comprising these statements first should be partitioned into relevant categories" (p.43). In the Glass study, as previously stated, the interview data are internally consistent, and we do not have access to the data from the field notes, onsite meetings and visits or document analysis. Rule 4, therefore, cannot be applied. Glass has woven information in her discussion that makes it reasonable to acknowledge consistency in those data, but the evidence instances are missing that would allow us to make a strong case.

We believe that the Glass study, in employing focused interviews using open ended questions and observations, makes a plausible case for qualitative confirmation of her hypothesis, but applying the M&F rules did not firmly settle the matter. The

study explores real life situations with no attempt to manipulate or control conditions. She argues that a high degree of autonomy is experienced by teachers and principals in both private and public schools, and that her findings disconfirm those of Chubb and Moe (1990) that teachers in private schools experience more autonomy than teachers in public schools. Regarding her second hypothesis, she identified six factors associated with autonomy: conflicting and contradictory demands, shared beliefs, layers of protection, a system of laws, funding constraints, and matters of the size of institutions. She concludes that autonomy is a complex process--an issue that does not distinguish the public from the private sector. As her only qualifier, Glass observes that similar organizational effects may not be encountered in schools under the duress of poverty and social dislocation, perhaps seeking to avoid the fallacy of confirming the consequent. It would appear from this example that qualitative confirmation yet eludes us; we revisit this in our appreciation.

Return to the Contents Table.

Go on to the next section (Appreciation).

An appreciation of the book by Miller and Fredericks, *Qualitative Research Methods*

Some of us attended graduate school and had our early experience in an educational research that was dominated by experimental psychology. Prompted by Campbell and Stanley (1963) we were concerned with the validity of our research, but we believed that if only we found the correct experimental design and carried it out competently our results would be valid. With methods based on the probability calculus, we could arrive confidently at a "significance level." Others of us know little of statistics and experimental design, having studied in programs and with faculty who do not believe in realism and seek, for example, verstehen-"a type of historical or contemporary insight which cognitively reconstructs a plausible interpretation of an action or event given knowledge of the cultural 'rules'." (M&F, p. 74) It may well be that the latter group is now dominant in educational research, and many are quite confident in their methods and believe their results adequately justified. Some believe no more justification is possible (e.g., Smith and Heshusius, 1986). At least a small subset of researchers remains uneasy that qualitative approaches lack means for validation. It is this group that M&F addresses, of course, those that would be extremely happy to have a means to attain "qualitative confirmation." In our opinion, graduate students (especially those considering qualitative methods for their research) should study and assess M&F's rules and rationale, whatever camp they are in. M&F's presentation makes this much more difficult than necessary, but the issue is important enough to make the effort.

The overall organization of chapters is logical; once apprehended it is clear:

1. The major issue in qualitative inquiry (confirmation-and introduction to rules)
2. Hypotheses in qualitative research methods (defense of hypotheses and the nature of evidence)
3. Additional rules of confirmation (M&F's version plus discussion of disconfirmation)
4. Assessing qualitative studies (applying the rules to some published reports)
5. Epistemological asides and conclusions (revisiting the intellectual roots)

Reading through the chapters, however, one finds complex concepts and intricate arguments that can only be clarified (usually!) by reading the appendixes:

1. Some notes on the nature of indeterminacy
2. Clarifying the "adequate evidence condition" in educational theory and research
3. Reciprocal paradigm shifts and educational research: A further view of the quantitative-qualitative dilemma.

After all this work, however, you will find that qualitative confirmation is still a judgment call. We attempted to illustrate this with our own application of the rules to the "Glass study". The published report of that study was attractive to us because the "evidence instances" were so obvious and because they constituted a "methodologically unique class". Because all the data (interviews) was available,

we could do some verification not usually available (search for negative instances not reported in the published version, for example). The studies analyzed by M&F in Chapter 4 did not provide as good a test of the method, but even this good test ended inconclusively, IOHO.

As with all works of genuine scholarship, one of the benefits of study is the acquaintance (or reacquaintance) it provides with key scholars and their ideas. M&F write from outside the "college" known best by the authors of this review, and we found many new and insightful sources. Most of us have heard of Quine but have not studied his books. Our inadequate preparation in philosophy left us ignorant of the seminal contributions of Winch, and we agree with M&F that, "It is amazing how much debate has been generated by the two rather modest works of Winch (1958, 1964)." (p. 74) Swinburne lurked unread in our library, as did Roth (shame, shame). We were led to an even stronger appreciation of Wolcott and introduced to the "cognitivist and semanticist", George Lakoff.

In summary, the presentation is sloppy, the writing dense and the organization suboptimal, but the topic is important and the fresh perspective welcome. We wish the authors would bring out an new and improved edition, but even if not we recommend the book to you.

Return to the Contents Table

References

Ball, S. J. (1987). *The micro-politics of the school*. New York, NY: Methuen and Company.

Blase, J. J. (1991). Some negative effects of principals' control-oriented and protective political behavior. *American Educational Research Journal, 27*, 727-753.

Campbell, D. & Stanley, J. (1963) Experimental and quasi-experimental designs for research on teaching. In, N. Gage (Ed.) *Handbook of research on teaching*. Chicago: McGraw-Hill.

Carnap, R. (1962). (2nd ed.). *Logical foundations of probability*. Chicago: University of Chicago Press.

Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools*. Washington, DC: The Brookings Institution.

Collin, F. (1985). *Theory and understanding: A critique of interpretive social science.* London: Basil Blackwell.

Coombs, C. (1964) *A theory of data*. New York: Wiley.

Davidson, D., (1984). *Inquiries into truth and interpretation*. Oxford: Oxford University Press.

Denzin, N.K. (1978). *The research act*. New York: McGraw - Hill.

Dreher, M. (1994). *Qualitative research methods from the reviewer's perspective, Critical Issues in Qualitative Research Methods.* J. Morse (Ed.), (pp.281-299). London: Sage.

Fuller, S. (1988) *Social epistemology*. Bloomington, IN: Indiana University Press.

Fogelin, R. (1982). (2nd ed.). *Understanding arguments*. Toronto: Harcourt Brace Jovanovich Inc.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.

Goetz J., & LeCompte, M. (1984). *Ethnography and qualitative design in educational research*. Orlando, Fl: Academic Press.

Greenwood, J. D. (1989) *Explanations and Experiment in Social Psychological Science: Realism and the Social Constitution of Action*. London: Springer-Verlag.

Guba, E.G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology Journal, 29*, 75-91, as cited in Owens, (1982). Methodological perspective: Methodological rigor in naturalistic inquiry: Some issues and answers. *Educational Administrative*

Quarterly, 18(2),1-21.

Hempel C. (1965). *Aspects of scientific explanation and other essays in the philosophy of science.* New York: Collier - MacMillan Limited.

Lakoff, G. (1987) *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

LeCompte, M.D., & Preissle, J. (1993). (2nd ed.). *Ethnography and qualitative design in educational research*. San Diego: Academic Press.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park: Sage.

Marshall C., & Rossman, G. (1989). *Designing qualitative research*. Newbury Park: Sage

Miles M.B., & Huberman, A.M. (1984*).* Driving valid meaning from qualitative data: Toward a shared craft. *Educational Researcher*, *13*(5), May 1984, 20-30.

Miles, M.B., & Huberman, A.M. (1984). *Qualitative data analysis*. Beverly Hills: Sage.

Miller, Stephen I. and Fredericks, Marcel (1994) *Qualitative Research Methods: Social Epistemology and Practical Inquiry*. New York: Peter Lang.

Moles, O. (Ed.) (1988). *High school and beyond administrator and teacher survey (1984): Data file users manual.* Washington, D.C.: Office of Educational Research and Improvement, U. S. Department of Education.

Owens, R.G. (1982). Methodological rigor in naturalistic inquiry: Some issues and answers. *Educational Administrative Quarterly, 18*(2), 1-21.

Polit, D. & B. Hungler (1995). *Nursing research principles and methods*, Philadelphia: J.B. Lippincott.

Popper, K. (1969). *The logic of scientific discovery*. New York: Science Education.

Quine, W. (1960). *Word and object*. Cambridge: Massachusetts Institute of Technology.

Quine, W. (1963). (2nd ed.). *From a logical point of view*. New York: Harper books.

Rosch, E. (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology: General, 104, 192-233.*

Roth, P., (1987) *Meaning and method in the social sciences: A case for methodological pluralism*. Itaca, NY: Cornell University Press.

Rosenberg, A. (1988). *Philosophy of social science*. Boulder, CO: Westview.

Scheffler, I. (1965) *Conditions of knowledge: An introduction to epistemology and education*. Glenview, IL: Scott, Foresman and Co.

Sedlak, M. W., Wheeler, C. W., Pullin, D. C. & Cusick, P. A. (1986). *Selling student short: Classroom bargains and academic reform in the American high school.* New York, NY: Teachers College Press.

Smith, J., & Heshusius, L. (1986) Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher, 15*, 4-12.

Spindler G., & Spindler, L. (1992). Cultural process and ethnography: An anthropological perspective. In M. LeCompte, W. Millroy, & J. Pressile (Eds.). *The Handbook of Qualitative Research in Education*. San Diego: Academic Press.

Stake, Robert E. (1995) *The Art of Case Study Research*. Thousand Oaks: Sage.

Strauss, A., & Corbin J. (1990*). Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park: Sage.

Swinburne, R. (1973). *An introduction to confirmation theory*. London: Methuen.

Talbot, L. (1995). *Principles and practices of nursing research*. St. Louis: Mosby.

Winch, P. (1964) Understanding a primitive society. *American Philosophical Quarterly, 1*, 307-324.

Winch, P. (1958) *The Idea of a Social Science and its Relation to Philosophy*. London: Routledge and Kegan Paul.

Wolcott, H.F. (1973*). The man in the principal's office: An ethnography*. New York: Holt, Rinehart and Winston.

Wolcott, H.F. (1990). *Writing up qualitative research.* London: Sage.

Wolcott, H.F. (1994). *Transforming qualitative data*. London: SAGE. Especially Chapter 11, On seeking-and rejecting validity in qualitative research, 337-373.

Return to:

- Purpose
- Description
- Related Works
- Rules
- Application
- Appreciation,
- Contents

Notes

(1) This is a collaborative review, so the authors are listed alphabetically.

Les McLean
lmclean@oise.utoronto.ca
Professor, Department of Curriculum, Teaching and Learning
Ontario Institute for Studies in Education/University of Toronto: OISE/UT
(416) 923-6641, ext. 2478

Margaret Myers
lmmyers@gtn.net
Doctoral Student, Department of Curriculum, Teaching and Learning
OISE/UT

Carol Smillie
csmillie@oise.utoronto.ca
Doctoral Student, Department of Curriculum, Teaching and Learning
OISE/UT

Dale Vaillancourt
dvaillancourt@oise.utoronto.ca
Doctoral Student, Department of Curriculum, Teaching and Learning
OISE/UT

(2)Quine uses the term "translation" in the usual sense of finding a representation in a language of a text expressed in another, broadening and deepening the discussion in *Word and Object* (1960). The term "translation manual" has been extended to describe how a researcher understands what people in a "foreign" culture say and do, how we make sense of field notes, for example. There will always be more than one possible translation manual for any situation, just as there is more than one possible translation of a text from English to French. That said, we may argue in favour of one particular translation manual, just as we may say we prefer one translation over another.

(3) "We will also use the phrase 'evidence-instance' to indicate that the qualitative data are now being recast as evidential statements for confirmation." *M&F*, p. 41.

(4) Recollected narrative-call it "personal communication".

(5) "obervation" (p. 8). "inherit" instead of "inherent" (p. 16). "accept" where "except" is intended (p. 22). "was" when "were" is correct (p. 44). view**ed** (p. 45)**.** "this not entail …" does need "does" (p. 94). "or" (not 'of') (p. 119). "intact" rather than "in tact" (p. 145). References are missing from the end of Chapter 3, e.g. Miller (1990)!

(6) M&F argue in the same spirit as Clyde Coombs (1964), who distinguished between "observations" and "data" (he dealt only with observations in the form of numbers). What M&F call data, Coombs called observations, which only became data after application of one of the scaling techniques coming to fruition and/or being developed by Coombs. The spirit is that information, whether qualitative or

quantitative, comes to every researcher first in a raw form that must be refined before it can be used to make inferences.

(7) Some of Hempel's rules (from M&F, p. 11):

9.1 Df. An observation report B directly confirms a hypothesis H if B entails the development of H for the class of those objects which are mentioned in B.

9.2 Df. An observation report B confirms a hypothesis H if H is entailed by a class of sentences each of which is directly confirmed by B.

9.3 Df. An observation report B confirms a hypothesis H if it confirms a denial of H.

9.4 Df. An observation report B is neutral with respect to a hypothesis H if B neither confirms nor disconfirms H.

(8) By a 'methodologically unique class', we mean a situation where the researcher employs one dominant form of data collection, such as interviews, for instance. While such a situation is probably not realistic, it is a logical possibility and for this reason is included. (p. 42).

Return to:

Purpose

Description

Related Works

Rules

Application of Rules

Appreciation