DOCUMENT RESUME

ED 318 080 EA 021 614

AUTHOR Marshall, Catherine
TITLE Goodness Criteria: Are They Objective Criteria or Judgment Calls?
PUB DATE 89
NOTE 26p.; Paper presented at the Alternative Paradigms Conference sponsored by Phi Delta Kappa (San Francisco, CA, March 25-26, 1989).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Bias; Educational Researchers; Elementary Secondary Education; *Evaluation Criteria; Higher Education; *Qualitative Research; *Research Problems; *Validity; *Values
IDENTIFIERS Empiricism; *Objectivity; Positivism; *Truth

ABSTRACT

In discussions about qualitative research, questions often arise concerning appropriate research focus, researcher objectivity, manipulation of people or environments examined, study completeness, and value considerations. Goodness questions (reliability, replicability, and validity) get mixed with value questions (generalizability), with frustrating results for many students. Research decisions depend on the tradition or paradigm framing the endeavor. Early empiricists accepted the following assumptions: (1) truth is knowable; (2) values can be excised from the research process; (3) good research is empowering; (4) research uncovers the researcher's and others' truths; and (5) research uncovers what people believe to be true. A sixth postpositivist version of assumption 1 states that truth has no absolutely authoritative foundation, but good research comes from good methodology properly applied. Good research is honest, open inquiry, where the researcher searches for alternative explanations and is self-critical. Evaluating the goodness and value of research requires a judgment call. Critical-scientist, constructivist, and postempiricist judges all have different evaluation criteria. In judging goodness criteria, qualitative researchers should look for 20 alternative paradigms. Those supporting postpositivist criteria must help colleagues construct new realities and resist being beaten into submission by quantitative researchers' skepticism. (24 references) (MLH)

Goodness Criteria: Are They Objective Realities or Judgment Calls?

Paper presented at the Alternative Paradigms Conference
San Francisco, March, 1989, by

Catherine Marshall, Vanderbilt University

GOODNESS CRITERIA:

ARE THEY OBJECTIVE REALITIES OR JUDGMENT CALLS?

When I first read the conference papers, the wonderful exchanges among Alice, the March Hare, and the Mad Hatter came to mind.

> "Then you should say what you mean," the March Hare went on.
>
> "I do," Alice hastily replied; "at least I mean what I say-- that's the same thing, you know."
>
> "Not the same thing a bit!" said the Hatter.
>
> "Why you might just as well say that 'I see what I eat' is the same thing as 'I eat what I see!'" (Carroll, 1915, pp. 100-101).

I often was as lost and impatient as Alice in the convoluted arguments about the nature of truth and goodness. Wouldn't it be more productive, I thought, for us to actually collect data and then argue about the persistent criteria of goodness questions. Suppose one-third went forth to study (Spradley, 1979) the hobos of San Francisco; one-third went forth to study the organization of work at the San Francisco Hilton; and the remainder studied the patterns of behavior at the annual meeting of the AERA. All would be doing field studies, in the natural setting, with an

exploratory stance. So all would inevitably confront the persistent goodness questions.

Persistent Goodness Questions

Questions that arise persistently in qualitative methods classes, doctoral committees and on editorial boards considering qualitative methodology are:

1. How do I decide which is the important research question or focus?

2. Must I become a part of the culture to truly understand it? If I do, how can I leave it and re-establish myself as an "objective" researcher?

3. Aren't I manipulating people and altering their environment when I enter it to study it?

4. How do I know how many cases or subjects or observations are enough?

5. How do I know when my study is complete?

6. How do I present my methods so that others will be convinced that what I found is the truth?

7. How do I prevent my values and biases from slanting the findings?

8. Could someone else do the same things I did and get the same findings?

9. How can I present my findings in such a way that they will be accepted and used by practitioners and policy makers?

Goodness questions (in ancient terminology: reliability, replicability, validity) are mixed with value questions (in ancient terminology generalizability). Students get angry, frustrated, when I so frequently reply, "it depends." Its a judgment call. It depends upon your purpose, your audience, and how finely tuned you are as a research instrument. But it also depends upon the tradition, as Jacob (1988) says, or on the paradigm framing your endeavor.

Assumptions About Goodness and Truth

Smith (1989) recounts the historical evolution of assumptions that affect goodness criteria. Early empiricists accepted the following:

Assumption #1: There is a knowable truth. The aim of research is to "see nature in its own terms" (Rorty, 1982, p. 192). This assumes that there is one knowable truth which can be discovered when we use proper methods well. Under this assumption, the criterion for goodness is the analysis of the goodness

of the use of the right method. We ask questions like "Was the t-test the right method?" or "Shouldn't he have used more participant observation and less focused interviewing?"

Assumption #2: Values can be excised from the research process. Here the criterion becomes the degree to which the researcher did separate "the facts" (a.k.a. findings) from the values (a.k.a. interpretation, recommendations, implications).

Empiricists did not worry about utility and contribution of research. They assumed that the scientists' work was complete when they did objective studies using proper methods. This would lead to law-like generalizations about human nature which would assist in shaping the efforts of those who would improve society.

But empiricism has failed to deliver on its promise to provide an intellectual practical mastery of the social world (Smith, 1989). With this failure came postempiricists' recognition that the theories and methods shape the knowledge obtained from research. So new rationales and value in research and corresponding new criteria of goodness were sought. From the

critical scientist represented by Popkewitz (1989), we find:

Assumption #3: Good research must empower people by helping them to see the historical meaning of events and to place themselves, their institutions, and their roles in historical context. Researchers should help people choose how they wish to respond and reshape their present and future. Thus good research includes an analysis of past truths.

The flaws in this assumption (i.e. historical "truths" are just as subjective and relative as current "truths") lead us to examine the next assumption:

Assumption #4: Research is a process in which the researcher uncovers his/her own as well as others' truths. Thus, good research must include a self-revelation. Only then can the researcher go forth and uncover the stories of others. This assumption leaves us again with no hope of discovering truth, we consider:

Assumption #5: Research is the process of uncovering what people believe to be true (regardless of any absolute truth). Thus, good research can be objective in that it uncovers what people believe--it

uncovers a variety of subjective truths. Goodness is judged by the degree to which the researcher explores the full range of beliefs and presents them clearly and objectively. Thus, a statement is true if it accurately depicts beliefs (Smith, 1989).

Finally we find a post-positivist version of Assumption #1:

Assumption #6: There is no such thing as an absolutely authoritative foundation of the truth but good research comes from good method properly applied. That is, good research is honest, open inquiry, where the researcher searches for alternative explanations and is self-critical. Thus, good research can be recognized by examining whether the researcher searched for alternative hypotheses, explored negative instances, examined biases, etc. Sloppy research is that which starts with an idea and goes forth in the real world to gather evidence supporting the idea, without being systematic in searching for the wide range of alternative explanation and versions of truth, and without trying to be self-critical.

Most post empiricists try to avoid relativism, the "anything goes" response. They acknowledge that, while

there is no one truth, there is a truth to be known about a range of beliefs at a given point in time. But there are disagreements even about that, as Smith (1989) demonstrates in his discussion of the disputes over Spiro and Leach's studies of The Tully River Blacks. Disputes arise over the meaning as well as the significance of findings. These disputes are often settled by examining who whether the studies were "properly done" or objective. This moves us back to discussion of what we mean by "properly done," a focus on procedures. It forces us to acknowledge again that every study has biases and presuppositions and that research involves the interpretation of the interpretations people give to their own situations (Smith, 1989). We are merely making a judgement when deciding whether one study is more valid than another. We have to decide whose biases were more correct. All of the earlier assumptions really lead to this reality: evaluating the goodness and value of research requires a judgment call.

Judges of Goodness

The metaphor of the courtroom serves to frame judgement calls. "Adverserial evaluation" used this metaphor: proponents of different evaluations of the same program must present their case and "duke it out"; the group with the best defense or offense is the group whose truth is to be viewed as valid. Extending this metaphor to the question of goodness in qualitative research, one then asks what counts as a good offense? a good defense? If it were a boxing match, what would win points? If it were a court of law, what would convince the judge or jury of the goodness of one piece of research?

Would a Critical Scientist judge, a Constructivist judge, or a Post-Empiricist judge have different criteria?

The conference keynote presenters imply each paradigm has its own assumption and goodness criterion.

For Post-Empiricist judges, the best case is the one that can document an honest, open, and careful procedure for arriving at a description of the beliefs that people hold.

Constructivist judges are happy with descriptions of the varied and multiple realities that are socially constructed. Good data are obtained by getting inside the worlds of others. The only truths are the emic realities of insiders. Certitude is not possible, there is no correct interpretation, and there is no end to the ever-evolving interaction that reveals meaning. In the Constructivist tradition, the debate over the goodness of one's interpretation is resolved by a dialogic process.

For Critical Theorist judges there are subjective and objective realities. Objective realities are the institutions, statuses, and roles created in history. The subjective are the individual's sensemaking processes. And neither the objective realities nor the subjective are free of the taint of prior social manipulation. So, good research is that which uncovers those manipulations, thus empowering people to see ways to control their own destinies. Social critique, to be good, however, must have an internal self-critiquing mechanism--the researcher must, for example, avoid the patronizing stance of findings that empower the research subjects to accept the researcher's view of

the best system. Otherwise the researcher simply becomes a new social controller.

Common Agreements on Criteria

Suppose, instead of abstractly debating goodness, we did go forth and conduct research on the hobos, the Hilton work norms, and the functions of convention behavior. We would each have varying foci, data collection strategies, different ways of using raw data to buttress our stories, and varying stances on whether to interfere with people and settings. How would we ever decide which studies were good? Which ones deserved publication? Which ones would speak to people in such a way that the human condition would be bettered? Which ones were most methodologically and ethically pure?

Would there be enough agreement on goodness to fend off those who would sanitize the research process and lay positivist criteria on qualitative research (Miles & Huberman, 1984). Would we have enough agreement to be able to denounce those try to make research researcher-proof, eliminating the biases but also the varying interpretation that come from the

human research instrument (see Marshall 1985b for a critique of this "testing-in-context" approach)? We would also shy away outright acceptance of "portraits" a la Lightfoot (1983). We would not accept apologies as in Boys in White (Becker, Greer, Hughes, & Strauss, 1961) saying "in one sense we had no research design" (p. 17).

I suggest that we would agree on common criteria (although different paradigms would weight each criterion differently). Most of us in judging the goodness qualitative studies would look for the following evidence: alternative paradigms. They are these:

1. The method is explicated in detail so the reader can judge whether it was adequate and makes sense. An articulate rationale for the use of qualitative methods is given so that skeptics will accept the approach. The methods for attaining entry and managing role, data collection, recording, analysis, ethics, and exit are discussed. There is an auditability trail--a running record of procedures (often done in an appendix), and there is description of how the site and sample were selected. Data

collection and analysis procedures are public, not magical.

2. Assumptions are stated. Biases are expressed, and the researcher does a kind of self-analysis for personal biases and a framework analysis for theoretical biases.

3. The research guards against value judgements in data collection and in analysis (i.e. avoiding transgression like Whyte's [1955] judgmental field notes about "dilapidated houses" in Street Corner Society).

4. There is abundant evidence from raw data to demonstrate the connection between the presented findings and the real world, and the data are presented in readable, accessible form, perhaps aided by graphics, models, charts, figures;

5. The research questions are stated, and the study answers those questions and generates further questions;

6. The relationship between this study and previous studies is explicit. Definitions of phenomena are provided, with explicit reference to previously

identified phenomena, but it is clear that the research goes beyond previously established frameworks--challenging old ways of thinking.

7. The study is reported in a manner that is accessible to other researchers, practitioners, and policymakers. It makes adequate translation of findings so that others will be able to use the findings in a timely way.

8. Evidence is presented showing that the researcher was tolerant of ambiguity, searched for alternative explanations, checked out negative instances, and used a variety of methods to check the findings (i.e. triangulation).

9. The report acknowledges the limitations of generalizability while assisting in the readers in seeing the transferability of findings.

10. It is clear that there was a phase of "first days in the field" where a problem focus was generated from observation, not from library research. In other words, it is a study that is an exploration, not merely a study to find contextual data to verify old theories.

11. Observations are made (or sampled) of full range of activities over a full cycle of activities.

12. Data are preserved and available for re-analysis.

13. Methods are devised for checking data quality (e.g. informants' knowledgeability, ulterior motives, and truthfulness) and for guarding against ethnocentric explanations.

14. In-field work analysis is documented.

15. Meaning is elicited from cross-cultural perspectives.

16. The researcher is careful about sensitivity of those being researched--ethical standards are maintained.

17. People in the research setting benefit in some way (ranging from getting a free meal or an hour of sympathetic listening to being empowered to throw off their chains).

18. Data collection strategies are the most adequate and efficient available. There is evidence that the researcher is a finely-tuned research instrument, whose personal talents, experiential biases and insights are used consciously. The researcher is careful to be self-analytical and recognize when she/he is getting subjective or going native.

19. The study is tied in to "the big picture." The researcher looks wholistically at the setting to understand linkages among systems.

20. The researcher traces the historical context to understand how institutions and roles have evolved. (For further discussion see Marshall, 1985a; 1985b)

For each of these criteria there are still judgment calls, cautions, limits, assumptions, and "buts." Smith (1989), and the conference keynote speakers Lincoln (1989), Popkewitz (1989), and Phillips (1989) are among those who have raised them. For example, how do you provide the historical context when history is socially constructed? Even if the researcher lays bare her/his theoretical and personal biases--what about the ones that are subconscious? What prevents the research from being simply explication of those biases, taking idiosyncratic, subjective researcher's view? When abundant thick description and raw data are presented, must the reader take it on faith that they are not selected to argue a position, prove a point, or find context-laden evidence of a previously existing theory or bias? On what basis does the researcher decide which evidence to present?

How can the study be reported in a manner accessible to practitioners and policy makers when they will only hear and use research that fits within traditional ways of thinking? The only "usable knowledge" is that which fits within policy makers' "assumptive worlds," (Lindblom and Cohen, 1979; Marshall, Mitchell & Wirt, 1989). How can one incorporate cross-cultural perspectives knowing that the practitioner does not want to read about how leadership in elementary schools compares to leadership tribes in Borneo? How can qualitative researchers show how <u>One Boy's Day</u> (Barker & Wright, 1951) or <u>The Man in the Principal's Office</u> (Wolcott, 1973) contains transferable, if not generalizable, knowledge, in spite of the small sample? Even when the researcher has explicitly detailed the procedures--one must still ask, do the methods not shape the findings?

Even with all of these questions, caveats, limits, and "buts," we do have criteria. Criteria of goodness, like our search for knowledge are ever-evolving. One judge puts more emphasis on criterion #1 than on criterion #21. Another counts #12 heavily, throwing

out studies lacking cross-cultural comparisons.
Judgment calls must still be made

Infiltrating to Construct New Realities

When editors, tenure and promotion committees, and policy makers judge qualitative research, they must know how to use post-positivist criteria, those appropriate for qualitative methods. We need to be on those boards and committees, armed with alternative criteria from alternative paradigms.

We also need to be articulate and passionate in advocating the importance of supporting researchers who go beyond the testing-in-context approach to explore for meaning. We need to build offensive weaponry to demonstrate the importance of post positivist research. We must show that:

1. Researchers cannot avoid the ethnocentrism of their field without methods that explore beyond dominant paradigms.

2. Many practices, policies, folk wisdoms, accepted myths, (e.g. rational bureaucracy, meritocracy) do not work, and so research methods are

needed to uncover hidden meanings, the subjective interpretations, the voices of the powerless.

3. Many dominant theories and assumptions were established inappropriately (e.g. administrative theory, decisionmaking theory, career development theory evolved without incorporating women's experience; assumptions about how to manage schools were developed in times of growth and with persistent myths about the separation of education and politics). Therefore, new grounded theory (Glaser & Strauss, 1967) must be developed by exploring how these theories fit with a wider reality.

4. There is great value in research on the dynamics of power, the social construction of reality, organizational myth-making and a focus on micro-politics and language, for, as Greenfield (1985) has said, "language is a dialect with its own army and navy" (p. 3).

5. Research can benefit from the intuitive and metaphorical insight of human researchers and "to isolate one psychological function, in this case the rational from all others, is to limit the effectiveness

of that function, diminish the quality of its product" (Barger & Duncan, 1982, p. 12).

6. When we allow those who are powerful to hold all the definitions of what is good and what is useful, we can never get beyond the dominant world views (Noblit, 1984).

Validity, practicality, utility, credibility, and worth are socially constructed judgment calls. We need to identify and build on our common agreements so that we too are participants in that social construction.

Too often, alternative paradigmers are beaten into submission as illustrated in this vignette, borrowed from Van Maanen (1979):

> Qualitative Researcher: "Many people these days are bored with their work and are...."
>
> Quantitative Researcher: "What people, how many, when do they feel this way, where do they work, what do they do, why are they bored, how long have they felt this way, what are their needs, when do they feel excited, where did they come from, what parts of their work bother them most, which...."

Qualitative Researcher: "Never mind" (p. 519).

REFERENCES

Bargar, R. R., & Duncan, J. K. (1982). Cultivating creative endeavor in doctoral research. Journal of Higher Education, (53)1, 1-31.

Barker, R. G., & Wright, H. F. (1951). One boy's day. New York: Harper.

Becker, Howard S., Geer, B., Hughes, E. C. and Strauss, A. L. Boys in White: Student Culture in Medical School. Chicago: University of Chicago Press, 1961.

Carroll, Lewis. (1915). Alice's adventures in wonderland. London: MacMillan.

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine.

Greenfield, T. B. (1985). Curriculum Inquiry, 15(1).

Lightfoot, S. L. (1983). The good high school. New York: Basic Books.

Jacob, E. (1988). Clarifying qualitative research: A focus on traditions. Educational Researcher, 16-24.

Lincoln, Y. (1989). The making of a constructivist: A remembrance of transformations past. Paper

presented for the International Conference on Alternative Paradigms. San Francisco, CA.

Lindblom, C. E., & Cohen, D. K. (1979). Useable knowledge: Social science and social problem solving. New Haven: Yale University Press.

Marshall, C. (1985a). Field studies and educational administration and policy: The fit, the challenge, the benefits and costs. Urban Education, 20, 61-80.

Marshall, C. (1985b). Appropriate criteria of trustworthiness and goodness in qualitative research on education or organizations. Quality and Quantity, 19, 353-373.

Marshall, C., Mitchell, D., & Wirt, F. (1986). The context of state level policy formulation. Educational Evaluation and Policy Analysis, 84, 347-378.

Marshall, C., Mitchell, D., & Wirt, F. (1989). Culture and education policy in the American states. Hampshire, England: Falmer.

Miles, M., & Huberman, A. M. (1984x). Qualitative data analysis: A sourcebook of new methods. Beverly Hills: Sage.

Noblit, G. W. (1984, Spring). The prospects of an applied ethnography of education: A sociology of knowledge interpretation. Educational Evaluation and Policy Analysis, 6(1), 95-101.

Phillips, D. C. (1989). Post-positivistic science: Myths and realities. Paper prepared for the International Conference on Alternative Paradigms, San Francisco, CA.

Popkewitz, T. S. (1989). Whose Future? Whose past? Notes on a critical science and methodology. Paper prepared for the International Conference on Alternative Paradigms, San Francisco, CA.

Rorty, R. (1982). Method, social science, and social hope. In R. Rorty, Consequences of pragmatism (pp. 191-210). Minneapolis: University of Minnesota Press.

Smith, J. K. (1989). Alternative research paradigms and the problem of criteria. Paper presented for the International Conference of Alternative Paradigms for Inquiry, San Francisco, CA.

Spradley, J. P. (1979). The ethnographic interview. New York: Holt, Rinehart & Winston.

VanMaanen, J., (Ed.) (1979). The seminar. Administrative Science Quarterly, 24(4), 519.

Whyte, W. F. (1955). Street corner society. University of Chicago Press.

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