

Criteria for Evaluation of Qualitative Research

Michèle Lamont
Department of Sociology
Harvard University

Draft Only—May 17, 2005

Comments prepared for the NSF Workshop on Qualitative Methods

A) Preliminary Remarks

I prepared my memo after most of the memos prepared for the workshop were posted. This was so that I could develop a sense of the center of gravity within our group, and work toward framing our task in a way that may be compatible with our general sense of what we want to do. Thus, you will see that my reflections were nourished by your own.

- 1) Debates about the evaluation of qualitative research within each of the fields represented at the workshop are contingent on the ecological environment in which they develop. In political science, much of the recent writings on the question have been framed in response to the challenges raised by King, Keohane and Verba's very influential *Designing Social Inquiry*. In anthropology, Clifford and Marcus's *Writing Culture* has generated a huge cycle of collective reflection on how the identity and the position of the researcher affects his work, which has framed the discussion about reliability in validity in terms almost incompatible with how the question is framed in psychometrics for instance (instead of bracketing the identity of the researcher, the challenge is to fully understand its impact on research). In sociology, a multimethod discipline par excellence, we are now going through a phase where more scholars are concerned with the similarities between the evaluation of qualitative and quantitative research, as a growing number of students are being trained for "multi-methods," at least in top departments. Yet, the repercussions of a long-lasting disciplinary split between quantoids and qualoids continue to be felt.
- 2) The impact of the broader disciplinary ecologies on the disciplinary conversations about the standards of qualitative work helps us understand the possibilities and limitations that our task may be facing. Because qualitative method is practiced in very different environments in anthropology, criminology/legal studies, political science, and sociology, we should focus less on our differences than on the **greatest common denominators shared across disciplines**. We have to focus on the perhaps somewhat small areas where standards of evaluation intersect. Furthermore, our charge is not to reflect on standards of evaluation for all qualitative research conducted in the social sciences, but only for the standards to be applied to proposals submitted to NSF. It goes without saying that many researchers may not want to conform to these standards.

- 3) My own insights into the topic at hand come from my experience as a qualitative sociologist, from teaching graduate seminars in qualitative research, and from conducting a research on definitions of quality used in funding panels in the social sciences and the humanities. This research focuses on how panelists draw the line between the winners and the losers, and on disciplinary differences in how scholars define quality across disciplines. The book *Cream Rising* will address the formal and informal criteria of evaluation used in interdisciplinary panels. As such, it takes as its object KNOWLEDGE EVALUATION PRACTICES. I consider the standards of evaluation of qualitative research to be a subset of this broader issue. Thus, at times, I will speak to what I think is right, or should be right, and at other times, I will speak to what I believe others believe is right.

B) Addressing the Questions

1) What are the standards of rigor in qualitative research?

The interviews I have conducted with panelists suggest that these standards vary depending on whether panelists adopt an epistemological standpoint influenced by a Weberian *Verstehen* approach to knowledge production (the most popular standpoint used by my respondents by far—see Figure 1), or by a positivist approach to knowledge production. Both standpoints are typically used by qualitative researchers. While the first group emphasizes the importance of insight, meaning-making, and attentiveness to context, the latter group emphasizes generalizability and, in some cases, falsification (Mallard, Lamont, and Guetzkow 2005). Nevertheless, in both cases, the fit between theory, methods, and data and the justification of the research procedures, including data analysis, are essential. In both cases, providing a cogent account of how various research decisions were made, and of how these tie with the theoretical motivation of the project, are a *sine qua non*. Thus, researchers of all stripes value making available to the evaluators the information needed to assess the work accomplished. Being explicit about how one goes about producing research, and thus allowing others to reproduce results if desired, is of the essence.

My study of funding panels, which in most cases considered proposals for qualitative research projects, also strongly suggests that most members of panels agree that they have shared standards of evaluation. This is confirmed by the fact that in most competitions, a consensus around a sizable proportion of the strongest and weakest proposals emerged in the rankings submitted by panelists prior to deliberation. This suggests that if panelists are not always able to articulate precisely what defines quality, many concur that “we recognize it when we see it.” But, what is this “it”? Attentiveness to details, clarity, and tightness of the connection between theory, data, and methods, are crucial, as is of course originality, defined largely as daringness, ability to frame a problem differently, or ability to open new vistas and ask new questions (on originality, see Guetzkow, Lamont, and Mallard 2004).

I noted above that in sociology, a growing number of departments abide by the motto of multimethod training, thereby sharing a common assumption that there is not one good

method, only good questions, and that the youngest generation should demonstrate an ability to mobilize the tools best suited to address a given question. This translates into what I perceive to be an important subcultural shift within the field: quantitative and qualitative researchers are moving away from thinking of themselves as distinct and even incompatible breeds. The best researchers are the ones who are able to do everything reasonably well. This shift also means that the standards of rigor in qualitative and quantitative research are increasingly convergent, as researchers use the same mental templates as they move across topics. However, qualitative research does not necessarily come to resemble more and more quantitative research (following a template proposed by King, Keohane, and Verba). Instead, there is growing awareness of the theoretical/empirical back and forth needed in qualitative and quantitative research alike, a perspective advocated for instance by Charles Ragin in his influential books. Instead of positing that quantitative research shows us the route for high-quality research, qualitative researchers are reflecting on the distinctive requirements to be met by qualitative research. In the study of racial inequality for instance, this translates into scholars becoming increasingly aware that the identity of the interviewer should not be bracketed, and that there are distinct advantages associated with having a group's insiders and outsiders interview someone, since the respondent's presentation of self is unavoidably shaped by whom she is talking to. The goal is not necessarily to produce generalizable knowledge, but to produce research that sheds new light on social phenomena and that adds specificity and complexity to our understanding. As proposed by David Snow (see memo for last NSF workshop), the goal may be providing theoretical refinement and extension to earlier formulations. These developments are entirely compatible with new trends in other fields—the new literature concerned with process tracing in political science, for instance.

Of course, I could go on to describe more specifically the standards of excellence in qualitative sociology, although this could be somewhat redundant given the excellent memos on this topic that were prepared for the last NSF workshop. Particularly useful in my opinion were the memo by Susan Silbey, which also focused on the importance of transparency and accountability, and again, the memo by Snow mentioned above. Again, an observation of evaluation practices suggests that shared standards exist. Beyond the context of my own research, qualitative sociologists are constantly asked to make collective decisions concerning academic products—dissertations, articles, fellowship applications. There are divergences concerning the details of the evaluation, but many organizational mechanisms are in place to encourage the development of a common discourse, to bracket alternative and incompatible standards, etc. In one of the papers that came out of our project, we show how panelists come to agree on the center of gravity of the group within the first few hours of a meeting, just as we did today as we tried to understand where others stood. Once this is accomplished, panelists adjust the formulation of their evaluations to make them somewhat palatable to and compatible with those of others, and to avoid open conflicts. Program officers encourage “translation,” i.e., they encourage panelists to formulate their arguments in a language that makes sense to others. They also have at their disposal a number of tools to avoid voting and to produce consensual decisions (Mallard, Lamont, and Guetzkow 2005).

2) How can the standards for rigor be communicated?

a) Within sociology

NSF could do much more to socialize applicants and panelists about what the common standards are. Several interesting alternatives were proposed in the memos prepared for this workshop—posting templates of good and bad projects, creating an institute for advanced studies in qualitative research, offering more discipline-wide teaching workshops, etc. I believe that the diffusion of more consistent standards will emerge as the number of funded qualitative proposals increases. I gather that NSF's interest in funding more qualitative research, and in standards for qualitative work, grew from an unanticipated and impressive increase in the number of qualitative dissertation improvement grant proposals in recent years. By writing these proposals and receiving feedback, graduate students are socialized into thinking more systematically about what defines a strong qualitative proposal. So do their mentors, many of whom have never sent a proposal to NSF (I know this from talking both with program officers and with top people in my field). As qualitative sociologists and political scientists, for instance, come to think that qualitative research of high quality *is* funded by NSF, they will put much more energy into producing top-notch proposals and into understanding what may define such proposals.

b) Across disciplines

Again, my research suggests that there is already quite a bit of overlap in standards across the social sciences, in that what we call the “constructivist standpoint,” which corresponds to Weber's *Verstehen*, seems to be privileged by the majority of social scientists whom I have interviewed for my project on funding panels. Of course, the consensus found among evaluators serving on competitions at the Social Science Research Council and the Woodrow Wilson Fellowship Foundation may be very different than those found at NSF, where panels tend to be uni-disciplinary and perhaps wedded to the more strongly institutional understandings of how each discipline differs from others. One could probably decide on a range of standards that are shared by social scientists across all the fields represented here, at least for the purpose of evaluation of NSF proposals. For instance, we would probably agree that proposals should be very explicit about the type of data the research will draw on and how the researcher will go about collecting the data. The areas of divergence—about, for instance, whether reflexivity is essential to good research—may better be left undiscussed. Again, a focus on the smallest common denominator can be much more easily reached than the adoption of a common set of principles that should be applied unequivocally to all fields (concerning generalizability or how to establish validity, for instance).

3) What are the most pressing issues of research and methods facing qualitative research projects?

We need to gather systematic information about how qualitative and quantitative research are similar and different from one another by interviewing researchers concerning how

they understand their work, e.g., whether they think knowledge “cumulates,” whether they produce “findings”(and what kinds of findings), and what they think gives value to various types of research, including specific recent examples of celebrated pieces of scholarship. The literature abounds with normative statements concerning the technical procedures to follow to produce high-quality research, but we know very little about the “laboratory life” of the social sciences. I believe that this is an essential next step. In particular we need to address very systematically how top-notch qualitative researchers understand their distinctive contributions to the social sciences. To take only one example, we need to explore what difference was made by Arlie Hochschild’s introduction of the concept of “emotion work” in gender relations (ditto for Ann Swidler and the cultural toolkit metaphor). Did these concepts allow for more knowledge cumulation and generalizability? If not, what *is* their contribution? How can we explain the appropriation of these tools by such a large number of researchers in recent years? What difference did they make to the knowledge we produce? I suspect that much of the qualitative research that has been most influential is research that has **generated new insights and opened new vistas**, not necessarily research that excelled if assessed by the standards of validity, replicability, and generalizability. This would certainly be the case for Geertz’s analysis of the Balinese cock fight, for Albert Hirshman’s *Voice, Exit, and Loyalty*, or Michael Walzer’s *Spheres of Justice*.

4) What areas or topics are most promising for investigation using qualitative methods?

I would not distinguish topics or areas along these lines. Most topics can be studied using qualitative or quantitative methods, and different aspects of a phenomenon are brought to light by the use of different techniques. Even our understanding of meaning making, a social process that is most frequently approached using qualitative, interpretive techniques, can be enriched by the use of more accurate measurement techniques (see, for instance, Jepperson and Swidler 1994; Mohr 1998).

5) What areas of promising qualitative research are most likely to foster multidisciplinary projects?

Because all social science topics are amenable to qualitative analysis, this is also a difficult question to answer. However, we can identify areas of growth and vitality within each discipline, areas toward which large numbers of young people are moving. Some of these areas or topics are particularly conducive to qualitative research. This is the case for cultural and economic sociology, or the study of social movements, for instance. One could imagine targeting such vital areas as starting points for building networks that would connect young people across a number of fields. The model for multidisciplinary work that is now being put in place by the Social Science and Humanities Research Council of Canada, building on the research program model of the Canadian Institute for Advanced Studies, is a case in point (see in particular the “successful societies program” at www.ciar.ca). These programs bring together researchers who share a research interest but have very different yet complementary expertise. The large “networks of excellence” put in place by the EU function on similar premises and have been extremely

successful, changing the face of European social science research (notably in the field of immigration; see Patricio 2004). They have been extremely successful at generating enormous dynamism across large multidisciplinary networks of scholars working on related topics.

6) What is needed to strengthen tools, training, data, research design, and infrastructure for conducting qualitative research?

Money and ideas. And, particularly acknowledging that qualitative research requires time, especially to immerse oneself in one's data so as to make sense of it.

7) What single advance in qualitative research method or approach would contribute most to strengthening qualitative research?

Acknowledging the multifaceted character of the social world, which will lead us to acknowledge the need for a diversity of research techniques and approaches. This in turn will lead to a greater appreciation of the specific contributions that can be made by the various approaches, and to recognizing that different standards should be used to evaluate different types of knowledge. Such standards should be conceived as complementary rather than exclusive of one another.

References

- Clifford, James, and George Marcus (ed.). 1986. *Writing Culture: The Politics and Poetics of Ethnography*. Berkeley, CA: University of California Press.
- Guetzkow, Joshua, Michele Lamont, and Gregoire Mallard. 2004. "What is Originality in the Social Sciences and the Humanities?" *American Sociological Review* 69 (2): 190-212.
- Ron Jepperson and Ann Swidler, 1994. "What Properties of Culture Should We Measure?" Meaning and Measurement in the *Sociology of Culture*, ed. Paul DiMaggio, special issue of *Poetics* 22: 359-71.
- King, Gary, Robert O. Keohane, and Sydney Verba. 1994. *Designing Social Inquiry. Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Lamont, Michele. Forthcoming. *Cream Rising: Defining Excellence in the Social Sciences and the Humanities*. Cambridge, MA: Harvard University Press.
- Mallard, Grégoire, Michele Lamont, and Joshua Guetzkow, 2005. "Epistemological Pluralism and Fairness in Peer Evaluation: Evidence from the Social Sciences and the Humanities." Under review.
- Mohr, John. W. 1998., "Measuring Meaning Structures." *Annual Review of Sociology* 24: 345-70.
- Patricio, Therese. 2004. *Towards an International Strategy with European Research Organizations* . Ottawa: Social Science and Humanities Research Council of Canada.
- Ragin, Charles C. 1994. *Constructing Social Research*. Thousand Oaks, CA: Pine Forge Press.
- Silbey, Susan, 2004. "Designing Qualitative Research Projects." Pp. 121-126 in *Workshop on Scientific Foundations of Qualitative Research*, Edited by Charles C. Ragin, Joane Nagel, and Patricia White. Arlington, VA: National Science Foundation.
- Snow, David A. 2004. "Thoughts on Alternative Pathways to Theoretical Developments: Theory Generation, Extension, and Refinement." P. 13 in *Workshop on Scientific Foundations of Qualitative Research*, Edited by Charles C. Ragin, Joane Nagel, and Patricia White. Arlington, VA: National Science Foundation.
- Weber, Max. [1905, 1991] 1949. *Max Weber on the Methodology of the Social Sciences*. Translated and edited by Edward A. Shils and Henry A. Finch. Shils. Glencoe, Ill: Free Press.

FIGURE 1
 THE DISTRIBUTION OF EPISTEMOLOGICAL STYLES USED BY EACH PANELIST BY
 DISCIPLINARY CLUSTERS AND DISCIPLINES

EPISTEMOLOGICAL STYLES	DISCIPLINARY CLUSTERS											Total				
	Humanities			History				Social Sciences								
	English	Musicology	Art History	Philosophy	History	History	Sociology	Political Sc.	Anthropo.	Economics	Geography		Science			
Constructivist	█					█						█				11
Comprehensive	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	38
Positivist						█	█			█					█	5
Utilitarian							█								█	14
Total	7	3	2	2		14		6		6		5	2	1	1	49

From Mallard, Lamont, and Guetzkow (2005).