

Rigor or rigor mortis: The problem of rigor in qualitative research revisited

Issues are raised by the persistent concern with achieving rigor in qualitative research, including the rigidity that often characterizes the search for validity in qualitative work and the threat to validity that the search for reliability may pose. Member validation is highlighted as a technique that exemplifies not only the practical, but also the profoundly theoretical, representational, and even moral problems raised by all procedures aimed at ensuring the trustworthiness of qualitative work. Key words: qualitative research, reliability, rigor, validity

Margarete Sandelowski, PhD, RN, FAAN
Professor
University of North Carolina at Chapel Hill
School of Nursing
Department of Women's and Children's Health
Chapel Hill, North Carolina

THE PROBLEM OF rigor in qualitative research continues to arouse, beguile, and misdirect. As researchers, we have a much clearer understanding of the challenges involved in producing good qualitative work and of techniques that can be used to ensure its trustworthiness. Yet we also remain in danger of succumbing to the "illusion of technique"^{1,2}: of making a fetish of it at the expense of perfecting a craft and of making rigor an unyielding end in itself. There is an inflexibility and an uncompromising harshness and rigidity implied in the term "rigor" that threaten to take us too far from the artfulness, versatility, and sensitivity to meaning and context that mark qualitative works of distinction. It is as if, in our quasi-militaristic zeal to neutralize bias and to defend our projects against threats to validity, we were more preoccupied with building fortifications against attack than with creating the evocative, true-to-life, and meaningful portraits, stories, and landscapes of human experience that constitute the best test of rigor in qualitative work.

In this article, I revisit the beguiling problem of rigor in qualitative research yet again.³ What seems clearer to me now—after many years of reading and doing qualitative work—and what I hope to clarify here, is that rigor is less about adherence to the letter of rules and procedures than it is about fidelity to the spirit of qualitative work.²

RECONCEPTUALIZING VALIDITY

Scholars have increasingly disputed the conflation of validity with either truth or value (a scientifically valid work may be neither true nor valuable), and the reification, commodification, and reduction of validity to a set of procedures. They have also increasingly abandoned the storybook image of science to focus on what scientists actually do. As Elliot Mishler concluded, they have come to recognize that the work of science is, in part, defined by the social process of validation.⁴

According to Mishler, different communities of researchers differently warrant and evaluate claims to scientific worthiness. Because no general rules can be provided for appraising validity in particular studies or domains of inquiry, and because no standard procedure can be determined either for assigning weights to different threats to validity or for comparing different kinds of validity, validation is less a technical problem than a deeply theoretical one. For Mishler, the evaluation of the trustworthiness of any single project is inevitably a matter of judgment, whereby skilled researchers use their tacit understanding of actual, situated practices in their fields of inquiry to do their own work, to make claims for it, and to evaluate the work of others. Mishler emphasized the social world

of scientists who strive to have their work accepted as good science. As social worlds are continually being made and remade, so are the practices that scientists use to support their claims-making. The social discourse on reliability, for example, is better understood as a particular way of warranting validity claims, rather than as a universal or abstract guarantor of truth. Indeed, as I discuss shortly, the effort to establish reliability (as it is conventionally portrayed in instructional literature) is often completely unwarranted in many qualitative projects and may, paradoxically, serve only to weaken claims to validity.

When validation is viewed as a culturally and historically situated social process, both experimentalist and interpretivist can be recognized as relying on contextually grounded linguistic and interpretive practices, rather than on rules assumed to be sufficiently abstract and universal for every project. Trustworthiness becomes a matter of persuasion whereby the scientist is viewed as having made those practices visible and, therefore, auditable; it is less a matter of claiming to be right about a phenomenon than of having practiced good science.

RELIABILITY AS A THREAT TO VALIDITY

One of the most important threats to the phenomenological validity⁵ (addressing the *true-to*) and, therefore, to the construct validity (addressing the *true about*) of qualitative projects is the assumption that validity rests on reliability. Investigators often claim that their findings are valid when, for example, they can show that research participants responded consistently over time and with each other concerning an experience,

In the naturalistic/interpretive paradigm, reality is assumed to be multiple and constructed rather than singular and tangible.

or that a panel of experts or persons other than the investigator coded information the same way. What is embedded in these examples is the notion of reality as external, consensual, corroboratory, and repeatable. What is being sought in these examples are coefficients of agreement or consensus on the nature of that reality.

What is forgotten is that in the naturalistic/interpretive paradigm, reality is assumed to be multiple and constructed rather than singular and tangible.⁶ Moreover, what is ignored in the indiscriminate transfer to interpretive research of the assumption that a valid work must be a reliable one is that qualitative research is an art, or, at least, as much art as science, and that the nature of the narrative data that are the mainstay of qualitative work is inherently revisionist. To put it in phenomenological terms, repeatability is not an essential (or necessary or sufficient) property of the things themselves (whether the thing is qualitative research or the qualitative interview).

Qualitative research as art

As described in detail elsewhere,^{7,8} there is a kinship between art and science, and qualitative research bridges these realms of meaning. Accordingly, any discussion of validity must occur in the context of the artfulness of qualitative inquiry. Renata Tesch elegantly summarized why the validity of an artful enterprise does not depend on repli-

cable outcomes.⁹ Indeed, according to Tesch, the result of a qualitative analysis should be viewed as

a representation in the same sense that an artist can, with a few strokes of the pen, create an image of a face that we would recognize if we saw the original in a crowd. The details are lacking, but a good “reduction” not only selects and emphasizes the essential features, it retains the vividness of the personality in the rendition of the face.^{9(p304)}

Similarly, a good qualitative data reduction grabs the “essence” of a phenomenon; it does not “flood (us) with so much detail (that we are) left with hardly a perception of the phenomenon at all.”^{9(p304)} Even when confronted with the same qualitative task, no two researchers will produce the same result; there will inevitably be differences in their philosophical and theoretical commitments and styles.

Tesch shows us not only why aiming for the so-called whole truth (“flooded with detail”) may actually interfere with apprehending truth (the “essence”), but also why scientific notions of replicability are often completely at odds with the phenomenological validity sought by researchers working within the naturalistic/interpretive paradigm. Just as Dali’s art is no less valid than Picasso’s by virtue of differently re-presenting common phenomena, so too may different qualitative re-presentations of common phenomena all be valid ones. “There is no one correct way of drawing a face.”^{9(p305)} The task for scholars in a practice-oriented discipline such as nursing is to find ways to apprehend and re-present these different representations to achieve the “fuller knowing” that advances knowledge and influences practice.

The revisionist nature of narratives

A second critical factor invalidating the assumption that a valid work is always a conventionally reliable one is the inherently revisionist nature of the stories participants tell us in interviews. As described in more detail elsewhere,¹⁰ these stories are remembrances about the past in a fleeting present moment soon to be past. Research participants often change their stories from one telling to the next as new experiences and the very act of telling itself cause them to see the nature and connection of the events in their lives differently. The idea of empirically validating the information in one story against the information in another for consistency is completely alien to the concept of narrative truth and to the temporality, liminality, and meaning-making function of stories.

The task for the researcher confronted with different versions of a life event is not to dismiss them as simply inconsistent with each other or to dismiss the storyteller as an unreliable informant. Rather, the researcher might, for example, consider whether the versions are truly inconsistent, or, if inconsistent, why discrepancies exist, or whether the two discrepant accounts even represent the same story. Although it may be a useful tool in some carefully selected instances, no reliability coefficient can ever adequately deal with the analytic challenges of narrative data.

MEMBER VALIDATION AS A THREAT TO VALIDITY

Member validation illustrates well not only the tenacity of the idea of reliability as the essential basis for validity, but also the

complexity of all such techniques directed toward ensuring the rigor of qualitative research. Member validation, or the member check, is a technique scholars have proposed for establishing the validity of researchers' interpretations of data collected from research participants and for ensuring that these participants have access to what has been made of their experiences. The member check, accordingly, involves a professional obligation to do good science and a specifically ethical obligation to support members' right to know. As various scholars have described it,^{6,11,12} member validation is an ongoing process throughout the life of a qualitative project. Researchers informally engage in member validation every time they seek clarification for or elaboration of meaning and intention from the people they interview or observe, or check out their evolving interpretations of the data they collect. Researchers formally engage in the process when they deliberately incorporate set procedures by which members can check the accuracy and adequacy of researchers' syntheses of data.

Member checking has been hailed as a way of enhancing the rigor of qualitative work by specifying a set of auditable practices and by virtue of its congruence with the qualitative goal of representing experience from the actor's point of view. Yet its potential to enhance qualitative work belies the deeply theoretical and ethical difficulties involved in this technique that may serve paradoxically to undermine the trustworthiness of a project. Indeed, what is often lost in the discourse on member checking is the recognition that both researchers and members are stakeholders in the research process, concerned with staking certain claims (to telling the truth, to being

right), with maintaining certain personas (as good persons, subjects, scientists), and with frequently divergent interests, commitments, and goals. Even when members and researchers seem to have the same goals (such as to tell a good story or to promote an agenda), they may not. After all, there are different stories to tell and different agendas to promote.

For example, members will inevitably look for themselves and their own reality in researchers' accounts of their lives, but researchers strive to represent multiple realities in a way that still remains faithful to each member's reality. I have found in my own work that members are sometimes more interested in concrete descriptions of their own experiences than in abstract syntheses that incorporate them with other members' experiences. I recall in one of my studies one member's comment—on a synthesis I had given her to check—that certain incidents never happened to her or that she had never felt certain feelings. Such comments suggest the investment members have in their own experiences and the difficulty they may have in recognizing other peoples' concrete experiences as variants of their own, or as part of a larger abstraction. Scientific abstractions may appear to the member to be far removed from the "conventionalities and literalnesses"^{13(p12)} of their own everyday lives. Indeed, "generalizations (of any kind) always tell a little lie in the service of a greater truth."^{14(p205)}

Both members and researchers are interested in accounts that represent experience fairly, but they may have very different views concerning what a fair account is. Whereas members may be motivated to consent to participate in research to justify their actions or to defend the inevitability of

certain outcomes, researchers may be motivated to conduct research to evaluate actions and to show the possibility of a variety of outcomes. Whereas members may strive to be accepted as good people, researchers typically strive to be accepted as good scholars; these goals may conflict.

The typically narrative nature of interview data makes the problem of determining accuracy of meaning and intention a deeply theoretical and moral one. As noted previously, the stories that members tell in interviews are themselves constantly changing. They represent members' efforts to order, find meaning in, and even live with the events in their lives at a particular moment in their lives. Stories previously told may elicit feelings members no longer have, regret, and/or have forgotten; a life event previously told as a tragedy may subsequently be told as a romance. Members may want such stories removed as data.

Again, the idea that information previously collected can be subsequently simply checked, corroborated, and/or corrected may be valid only for certain cases. Stories are not simply vehicles for the communication of information that can be easily categorized and counted for consistency; rather, they are time-bound interpretive, political, and moral acts. Researchers must account for both the informational contents in and the discursive features of interview data: for what is said and meant in the interview and for how the interview itself was made. Researchers employing the member-checking process are always obligated to ensure that any correction of contents or feeling tone is warranted as a correction and not as a new story that must be analyzed for its meaning and relation to other stories. Researchers may make serious analytic errors

in attempting to find temporal, informational, or intentional consistency among stories. Moreover, analytic decisions become moral ones in the case of participants who wish to retract or alter information previously provided. Significantly, the information from any one interview "cannot be simply decontextualized to constitute a test of validity."^{11(p164)}

Members may also simply not be in the best position to check the accuracy of an account. They may have forgotten the information they provided or the manner in which it was provided. A member in one of my studies insisted that a small portion of a transcript of an interview was incorrect. Yet, both the transcriptionist and I, while proofing the transcript against the audio-taped interview, heard her words as they appeared in the transcript. What the woman forgot was that she was crying during this portion of the interview and likely made statements she now sees as wrong. What I had to decide was whether the disputed information was analytically important; that is to say, did I have to resolve the problem of accuracy here? If I had to resolve it, how would I do it? For example, if I played the tape back to the woman, she would be forced to hear herself at a time when she was very distressed, and she would also be proven wrong about what she thought she had said. How would this implementation of the member-checking process affect her

emotional health? These are only a few of the practical, theoretical, and ethical questions raised by this one incident alone. (As it turned out, the disputed lines were theoretically unimportant. I simply thanked her for correcting the lines and made a note of it on the transcript.)

Interestingly, a few participants have asked me for copies of their transcripts as remembrances for themselves. I provide these transcripts with the caveat that it may be difficult for them to relive the moment in time captured in those transcripts. For members, the effect of seeing in print what they once said or listening to themselves on tape may be similar to the effect of seeing oneself on videotape giving birth: somewhat bizarre and not wholly comfortable.¹⁵ Importantly, whether shared with members as a remembrance or as part of a validation process, such play-by-play written or taped accounts of past events have effects that researchers have yet fully to explore.

Researchers engaged in member checking also have the problem of determining when to initiate a formal member-checking process and what synthesis of data to present to members to check. Because the member-checking process is itself a variable that may influence the findings, researchers have to make decisions about when during the research project to initiate formal procedures. The very act of reading a transcript for accuracy may not only lead the member to provide additional data that have to be analyzed, but it may itself also cause the member to revise his or her views and/or influence events still to be experienced in the course of the study.

Researchers may offer the member some lay rendition of the findings written or presented in everyday language accessible to

Researchers may make serious analytic errors in attempting to find temporal, informational, or intentional consistency among stories.

the general public, or a scholarly synthesis that may be accessible only to other scholars. I recall one member who agreed to my everyday-English synthesis of a portion of her life, but who had difficulty understanding the scholarly synthesis prepared for a professional journal. The target audience for our work is often not a member group, but rather our own peers. Lay and scholarly syntheses are necessarily different from each other. Because they must adhere to different rules for representing data and often reflect different purposes, these syntheses may not be consistent with each other. Committed as many of us are to being accepted as good scholars and to telling scientific truths, we may say one thing in a report the member sees and another in a report the member does not even comprehend or may never see; neither of these versions will be lies nor will they constitute the whole truth. Importantly, the problem of member checking confronts researchers with all of the representational problems involved in trying to appeal to different audiences and in choosing the point of view and voice that will prevail in our reports of research. Indeed, we have yet to appreciate fully that decisions about how to present findings necessarily involve moral choices and that the conventional research report may be neither faithful to the letter or spirit of the phenomena addressed nor morally justified. In disseminating the findings of our research, we have yet to consider seriously the ethical implications of the choices we make or how artists (fiction writers, painters) resolve dilemmas involving differing voices and points of view.

Finally, member checking is itself socially constructed by the artifices and conventions of social interaction and research.

Accordingly, members may participate in a formal checking process only to meet the expectations of researchers and to be good subjects; members may be uninterested in participating in such an exercise. In order to minimize conflict, members may be reluctant to disagree with researchers' interpretations. Moreover, different members may have very different views of the same interpretation. Researchers have to address how the artifice of the research process itself may influence the validation process and whether the lack of convergence or consensus between researcher and member or among members themselves necessarily invalidates an interpretation.

As Bloor summarized it, members' responses are not "immaculately produced,"^{11(p171)} but rather they are shaped and constrained by such factors as the: (1) nature of the interaction between researcher and member, member and member, and among researcher, member, and the audiences to which they desire to appeal; (2) social norms concerning politeness and consensus building; and, (3) frank conflicts of interest and need. Despite our claims to be "doing what comes naturally,"^{6(p187)} the research process is inherently social and, therefore, itself subject to analysis as data in our studies.



Member checking exemplifies the practical and deeply theoretical, representational, and even moral problems involved in using such techniques. Indeed, practical problems are frequently theoretical, and representational problems are frequently both theoretical and moral. Although such techniques hold the promise of making the practices of qualitative researchers more visible and ac-

ceptable as science, they may cause as many problems as they resolve. Similar problems exist with such validation strategies as the expert panel, peer debriefing, and triangulation, where consensus or convergence may also be inappropriately sought, where the wrong expertise may be sought (for example, analytic expertise may be more relevant to a validation enterprise than clinical expertise), or where peers are also motivated by certain interactional constraints. We shall succeed in our efforts to ensure trustworthiness to the extent that we recognize the complexity of these strategies, analyze them critically, and select among them carefully. We shall succeed in our efforts to

the extent that we are willing to consider also artistic resolutions to such problems as whose voices will be heard and what narrative stance to assume in reports of research.

Research is both a creative and a destructive process; we make things up and out of our data, but we often inadvertently kill the thing we want to understand in the process. Similarly, we can preserve or kill the spirit of qualitative work; we can soften our notion of rigor to include the playfulness, soulfulness, imagination, and technique we associate with more artistic endeavors, or we can further harden it by the uncritical application of rules. The choice is ours: rigor or rigor mortis.

REFERENCES

1. Barrett W. *The Illusion of Technique*. New York, NY: Anchor Books; 1978.
2. Van Manen M. *Researching Lived Experience: Human Science For An Action Sensitive Pedagogy*. Albany, NY: State University of New York Press; 1990.
3. Sandelowski M. The problem of rigor in qualitative research. *ANS*. 1986;8:27-37.
4. Mishler EG. Validation in inquiry-guided research: the role of exemplars in narrative studies. *Harv Educ Rev*. 1990;60:415-442.
5. Bronfenbrenner U. The experimental ecology of education. *Teachers Coll Rec*. 1976;78:157-204.
6. Lincoln YS, Guba E. *Naturalistic Inquiry*. Beverly Hills, Calif: Sage; 1985.
7. Sandelowski M. The proof is in the pottery: toward a poetic for qualitative inquiry. In: Morse JM, ed. *Critical Issues in Qualitative Research*. Newbury Park, Calif: Sage; In press.
8. Sandelowski M. *Truth/Story-Telling in Nursing Inquiry*. Paper presented at Institute for Philosophical Nursing Research, Third Invitational Conference on Philosophy in the Nurse's World; May 4, 1993; University of Alberta, Banff, Alberta, Canada.
9. Tesch R. *Qualitative Research: Analysis Types and Software Tools*. New York, NY: Falmer Press; 1990.
10. Sandelowski M. Telling stories: narrative approaches in qualitative research. *Image J Nurs Sch*. 1991;23:161-166.
11. Bloor MJ. Notes on member validation. In: Emerson RM, ed. *Contemporary Field Research: A Collection of Readings*. Boston, Mass: Little, Brown; 1983.
12. Hoffart N. A member check procedure to enhance rigor in naturalistic research. *West J Nurs Res*. 1991;13:522-534.
13. Nisbet R. *Sociology As An Art Form*. New York, NY: Oxford University Press; 1976.
14. Barley N. *Not A Hazardous Sport*. New York, NY: Henry Holt; 1988.
15. McKay S, Barrows TL. Reliving birth: maternal responses to viewing videotape of their second stage labors. *Image J Nurs Sch*. 1992;24:27-32.