
Article

Reading Qualitative Studies

Margarete Sandelowski
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina, USA

Julie Barroso
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina, USA

© 2002 Sandelowski et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

In this article, the authors hope to shift the debate in the practice disciplines concerning quality in qualitative research from a preoccupation with epistemic criteria toward consideration of aesthetic and rhetorical concerns. They see epistemic criteria as inevitably including aesthetic and rhetorical concerns. The authors argue here for a reconceptualization of the research report as a literary technology that mediates between researcher/writer and reviewer/reader. The evaluation of these reports should thus be treated as occasions in which readers seek to make texts meaningful, rather than for the rigid application of standards and criteria. The authors draw from reader-response theories, literature on rhetoric and representation in science, and findings from an on-going methodological research project involving the appraisal of a set of qualitative studies.

Keywords: Reader-response, reading, representation, rhetoric, qualitative research, quality criteria, writing

Acknowledgments: We thank the members of our Expert Panel (named in Note 2) and our research assistants, Janet Meynell and Patricia Pearce.

Introduction

Over the past 20 years, reams of articles and books have been written on the subject of quality in qualitative research. Addressing such concepts as reliability and rigor, value and validity, and criteria and credibility, scholars across the practice and social science disciplines have sought to define what a good, valid, and/or trustworthy qualitative study is, to chart the history of and to categorize efforts to accomplish such a definition, and to describe and codify techniques for both ensuring and recognizing good studies (e.g., Devers, 1999; Emden & Sandelowski, 1998, 1999; Engel & Kuzel, 1992; Maxwell, 1992; Seale, 1999; Sparkes, 2001; Whittemore, Chase, & Mandle, 2001). Yet after all of this effort, we seem to be no closer to establishing a

consensus on quality criteria, or even on whether it is appropriate to try to establish such a consensus. Garratt and Hodkinson (1998) questioned whether there could ever be "preordained" (p. 517) "criteria for selecting research criteria" (p. 515). Sparkes (2001) stated it was a "myth" that qualitative health researchers will ever agree about validity. And Kvale (1995) suggested that the quest for quality might itself be an obsession interfering with quality.

The major reason for this lack of consensus is that no "in principle" (Engel & Kuzel, 1992, p. 506) arguments can be made that can uniformly address quality in the varieties of practices designated as qualitative research. As Schwandt (2000, p. 190) observed, qualitative research is "home" for a wide variety of scholars across the disciplines who appear to share very little except their general distaste for and distrust of "mainstream" research. Indeed, these scholars are often seriously at odds with each other. Accordingly, it is not surprising that these different communities of qualitative researchers have emphasized different quality criteria. Standards for qualitative research have variously emphasized literary and scientific criteria, methodological rigor and conformity, the real world significance of the questions asked, the practical value of the findings, and the extent of involvement with, and personal benefit to, research participants (e.g., Emden & Sandelowski, 1998, 1999; Heron, 1996; Lincoln & Reason, 1996; Richardson, 2000a,b; Whittemore, Chase, & Mandle, 2001).

Moreover, we have found in our own work that even when ostensibly the same criteria are used, there is no guarantee that reviewers will use them the same way, agree on whether a study has met them, or, if they agree, have the same reasons for agreeing. Indeed, we recognized, from our own efforts to review approximately 70 published reports and dissertations on women with HIV/AIDS, how little we consistently relied on any one set of criteria for evaluating qualitative studies, but how much we relied on our own personal readings and even "re-writings" of the reports themselves. While one of us tends to assume an "aesthetic" stance toward research reports, responding in terms of her total engagement with texts, the other tends to assume an "efferent" stance, reading primarily for the clinically relevant information they provide (Rosenblatt, 1978).

In this article, we hope to shift the debate in the health-related practice disciplines concerning quality in qualitative research from a preoccupation with epistemic criteria toward consideration of aesthetic and rhetorical concerns. Indeed, we see epistemic criteria as inevitably including aesthetic and rhetorical concerns. We argue here for a reconceptualization of the research report as a dynamic vehicle that mediates between researcher/writer and reviewer/reader, rather than as a factual account of events after the fact. More specifically, we propose that the research report is more usefully treated as a "literary technology" (Shapin, 1984, p. 490) designed to persuade readers of the merits of a study than as a mirror reflection of that study. We further propose that the evaluation of these reports be treated as occasions in which readers seek to "make meaning" from texts (Beach, 1993, p. 1), rather than for the rigid application of standards and criteria.

To make our case, we draw from reader-response theories (Beach, 1993) which emphasize the interactions between readers and texts by which "virtual texts" (Ayres & Poirier, 1996) are produced, and from studies of rhetoric and representation in science and ethnography, which emphasize the writing practices intended to produce appealing texts (e.g., Clifford & Marcus, 1986; Geertz, 1988; Hunter, 1990; Latour & Woolgar, 1986; Lynch & Woolgar, 1988). We describe key problems in using existing guides for evaluating qualitative studies and offer a reading guide that addresses these problems.

The Qualitative Metasynthesis Project

We draw also from the findings of an on-going research project, the purpose of which is to develop a comprehensive, usable, and communicable protocol for conducting qualitative metasyntheses of health-related studies, with studies of women with HIV/AIDS as the method case.¹ We began the project in June of 2000 and we expect to complete it by June of 2005. The reading guide we feature here is a product of that study, a key component of which is the development of a tool that will allow the systematic appraisal of a set of qualitative studies but also the preservation of the uniqueness and integrity of each individual study in the set. In order to enhance the validity of the findings of this project, we chose an expert panel comprised of six scholars who have had experience conducting qualitative metasynthesis and/or quantitative meta-analysis projects.² Their role is to provide peer review of our procedures: that is, to assist us to "think/talk aloud" (Fonteyn, Kuipers, & Grobe, 1993) the often inchoate moves that comprise so much of qualitative work. They will also evaluate the contents, relevance, and usability of the various tools — and the protocols for using them — we will develop for conducting each phase of a qualitative metasynthesis project, from conceiving such a project to disseminating the findings from it to various audiences.

We convened the members of the expert panel³ in March of 2001 for a two-day discussion (the contents of which were transcribed) of the version of the guide we had at that time, which was based on our intensive analysis of those reports of qualitative studies on women with HIV/AIDS we had retrieved by then. In order to prepare for this meeting, the panel members used the guide with a study we had purposefully selected from our bibliographic sample as an example of what we then thought was a methodologically "weak" albeit informative study. After this meeting, we revised the guide again, used it on all of the studies we had retrieved to date, and then asked the expert panel members⁴ to use this latest version with a purposefully selected sample of five other studies from our bibliographic sample. This time, we chose studies to represent variations in medium (e.g., journal article, book chapter), complexity (Kearney, 2001) (e.g., low-complexity descriptive summaries to high-complexity grounded theories and cultural interpretations), style of presentation (e.g., traditional scientific, alternative), and in author affiliations (e.g., nurse, social scientist). We asked the panel members to use the guide with — and to comment on its contents and usability for — each of these five studies and then to rank the relevance to them for each study of each category of information listed in the guide. We refer to the results of this work later in this article.

Reading and Writing Qualitative Studies⁵

Readers of research reports bring to these texts a dynamic and unique configuration of experiences, knowledge, personality traits, and sociocultural orientations. Readers belong to one or more "interpretive communities" (Fish, 1980) (e.g., qualitative researchers, academic nurses, social constructionists) that strongly influence how they read, why they read, and what they read into any one text. The members of these communities differ in their access and attunement to, knowledge and acceptance of, and participation with, for example, references and allusions in a text, the varied uses of words and numbers, and various genres or conventions of writing. Because of their varying reading backgrounds, experiences, and expectations, readers will vary in their interaction with texts (Beach, 1993; Lye, 1996a,b). Indeed, even when one reader is engaged with the same text, interactions will vary as such factors as the passage of time and different reasons for reading that text alter the reading. Moreover, reading is cumulative as each new reading builds upon preceding readings of this

and other texts (Manguel, 1996).

Researchers/writers, in turn, employ various writing conventions and literary devices in order to appeal to readers, and to shape and control their readings. Shape is a property of information that includes, not just the informational content per se, but also the very physical form in which it appears (Dillon & Vaughan, 1997). Indeed, the research report is itself better viewed, not as an end-stage write-up, but rather as a dynamic "literary technology" (Shapin, 1984) whereby writers use literary devices — such as correlation coefficients, p values, metaphors, and coding schemas — rhetorically to engage readers to accept their study procedures and findings as valid. As Shapin (1984, p. 491) conceived it, this technology is intended to make readers "virtual witness(es)" to what they have never seen: namely, the conduct of the project itself. Researchers/writers "deploy...linguistic resources" (Shapin, 1984, p. 491), such as the correlation coefficient and the emotive quote, to appeal to communities of scholars that will find such appeals convincing. In the case of the correlation coefficient, the appeal is to stability and consensus; in the case of the quote, the appeal is to "giving voice." These devices contribute to the illusion that write-ups of research are reflections of reality and that readers are witnessing the study reported there. Graphs, charts, tables, lists, and other such visual displays are powerful rhetorical devices that function "manifestly" to reduce large quantities of data into forms that can be more readily apprehended by readers, but also "latently" to shape findings and to persuade readers of the validity of findings (McGill, 1990, p. 141). They are components of the literary technology of science, not only because they evoke images of the research that has taken place, but also because they themselves constitute a visual source of information. They are part of the "iconography" of science, offering "visual assistance" to the virtual witness (Shapin, 1984, pp. 491-492).

In a similar rhetorical and representational vein, statistics are not merely numeric transformations of data, but "literary...displays treated as dramatic presentations to a scientific community" (Gephart, 1988, p. 47). In quantitative research especially, the appeal to numbers gives studies their rhetorical power. Statistics are a naturalized and rule-governed means of producing what is perceived to be the most conclusive knowledge about a target phenomenon (John, 1992). John (1992, p. 146) proposed that statistics confer the "epistemic authority" of science. The power of statistics lies as much in their ability to engender a "sense of conviction" (John, 1992, p. 147) in their "evidentiary value" (p. 144) as to provide actual evidence about a target phenomenon. Statistics authorize studies as scientific and contribute to the "fixation of belief" whereby readers accept findings as facts and not artifacts (Amann & Knorr-Cetina, 1988, p. 85). They are a display of evidence in the "artful literary display" (Gephart, 1988, p. 63) we know as the scientific report, and they are a means to create meaning. Statistical meaning is not "inherent in numbers," but rather "accomplished by terms used to describe and interpret numbers" (Gephart, 1988, p. 60). Indeed, quantitative significance is arguably less found than created, as writers rhetorically enlist readers, with the use of words such as *high* and *substantial*, to accept their findings as significant. Writers do not find, so much as they participate with willing readers to create, quantitative significance (Gephart, 1986). Communities of scholars are created and sustained by virtue of this participation (Clark, 1990).

Whereas tables and figures provide much of the appeal in quantitative research, tableaux of experience and figures of speech provide much of the appeal in qualitative research. Writers wanting to write appealing qualitative research reports tend to use devices, such as expressive language, quotes, and case descriptions, in order to communicate that they have recognized and managed well the tensions, paradoxes, and contradictions of qualitative inquiry. Qualitative writers desire to tell "tales of the field" (Van Maanen, 1988) that convey methodological rigor,

but also methodological flexibility; their ability to achieve intimacy with, but also to maintain their distance from, their subjects and data; and, their fidelity to the tenets of objective inquiry, but also their feeling for the persons and events they observed. They want their reports to be as true as science is commonly held to be, and yet as evocative as art is supposed to be.

In summary, the only site for evaluating research studies — whether they are qualitative or quantitative — is the report itself. The "production of knowledge" cannot be separated from the "communication of knowledge" by which "communities" of responsive readers are created and then come to accept a study as valid (Shapin, 1984, p. 481). The production of convincing studies lies in how well the needs and expectations of readers representing a variety of interpretive communities have been met. Indeed, although we tend to distinguish between epistemic and aesthetic criteria, they are in practice indistinguishable as the sense of rightness and feeling of comfort readers experience reading the report of a study constitute the very judgments they make about the validity or trustworthiness of the study itself (Eisner, 1985). As Eisner (1985) observed, all forms — whether novel, pottery, or scientific report — are evaluated by the same aesthetic criteria, including coherence, attractiveness, and economy. Quantification and graphical displays are common ways to achieve these goals in science texts (Law & Whittaker, 1988), while conceptual renderings, quotes, and narratives are common ways to achieve these goals in ethnographic texts. The aesthetic is itself a "mode of knowing" (Eisner, 1985); both scientific and artistic forms are judged by how well they confer order and stimulate the senses. Whether a reviewer judges study findings as vivid or lifeless, coherent or confusing, novel or pedestrian, or as ringing true or false, s/he is ultimately making a communal, but also a personal (Bochner, 2000; Richardson, 2000a) and an "aesthetic judgment" (Lynch & Edgerton, 1988, p. 185).

The Problem With Existing Guides for Evaluating Qualitative Studies

Although useful, existing guides for evaluating qualitative studies (variously comprised of checklists and/or narrative summaries of criteria or standards) tend to confuse the research report with the research it represents. They also do not ask the reviewer to differentiate between understanding the nature of a study-as-reported and estimating the value of a study-as-reported, nor do they allow that any one criterion might be more or less relevant for any one study and to any one reviewer. We prefer the word *appraisal* as opposed to evaluation, as appraisal more explicitly encompasses understanding in addition to estimating value. Any work of art — including the research report — must be understood, or appreciated, for what it is before it can be judged as a good or bad example of its kind. Appreciation here means the exercise of wise judgment and keen insight in recognizing the nature and merits of a work. Whether in science or art, scholarly criticism entails connoisseurship, or readers with the ability to appreciate the range and diversity of qualitative research (Eisner, 1991). If writers of qualitative research reports have an obligation to write well, readers of these reports have an obligation to read well.

Form is Content

Because the only access any reader/reviewer typically has to any study is to a report of it, what a reader/reviewer is actually reading/appraising is not the study itself, but its representation in some publication venue, usually a professional journal. The research report is an after-the-fact reconstruction of a research study and generally one that makes the inquiry process appear more orderly and efficient than it really was. In the health sciences, reports of research generally conform to what Bazerman (1988) described as the experimental scientific report. This literary style of reporting research is a "prescriptive rhetoric" (Bazerman, 1988, p. 275) for reporting

research that conceives the write-up as an objective description of a clearly defined and sequentially arranged process of inquiry, beginning with the identification of a research problem, and research questions or hypotheses, progressing through the selection of a sample and the collection of data, and ending with the analysis and interpretation of those data (Golden-Biddle & Lock, 1993; Gusfield, 1976).

The standardization of form evident in the familiar experimental scientific report does not so much reflect the procedures of any particular study as it reinforces and reproduces the realist ideals and objectivist values associated with neo-positivist inquiry. Written in the third person passive voice; separating problem and questions from method, method from findings, and findings from interpretation; and, representing inquiry as occurring in a linear process and findings as truths that anyone following the same procedures will also find, these texts reproduce the neo-positivist assumption of an external reality apprehendable, demonstrable, and replicable by objective inquiry procedures. The reader/reviewer knows what to expect in the conventional science write-up, and the fulfillment of this expectation alone constitutes a major criterion by which s/he will evaluate the merits of study findings. A write-up that fails to meet reader/ reviewer expectations for the write-up will jeopardize the scientific status of the study it represents (McGill, 1990). Although standardization of form is actively sought in the belief that form ought not to confound content, form is inescapably content. Researchers/writers are expected to report their studies *as if* the in vivo execution of these studies conformed to the prescribed form for reporting them. They are expected to make real life conform to the page.

Complicating the critique (that is, appreciation + appraisal) of qualitative reports is that many of them do not conform to the conventional experimental style. A hallmark of qualitative research is "variability," not "standardization" (Popay, Rogers, & Williams, 1998, p. 346), including the reporting of findings. Many qualitative researchers do not adhere to neo-positivist tenets and thus seek to write in ways that are more consistent with their beliefs. Moreover, it is a commonplace in qualitative research that "one narrative size does not fit all" (Tierney, 1995, p. 389) in the matter of reporting qualitative studies. Indeed, there is a burgeoning effort — itself a target of criticism (e.g., Schwalbe, 1995) — to experiment with different forms for communicating the findings of qualitative studies, including novels, poems, drama, and dance (Norris, 1997; Richardson, 2000b).

Accordingly, in order to appraise a qualitative study fairly, readers/reviewers have to appreciate the various forms and "narrative sizes" that qualitative reports come in so that they will know what they are looking at, what to look for, and where to find it. For example, we found in our review of studies of women with HIV/AIDS that some reports had no explicit description of method either in sections devoted just to this topic or anywhere else in the report, nor any explicit statement of research questions. Yet method was still discernible in the findings. Some reports had no explicit statement of a problem, which was instead implied in the research purpose and/or literature review, or became evident in the findings. Although some readers will not mind having to read method into the findings or the problem into the literature review, other readers will insist that writers explicitly address these categories of information.

Reporting Adequacy Versus Procedural or Interpretive Appropriateness

Another problem with existing guides for evaluating qualitative studies is that they do not clearly separate for the reader what a writer reported that s/he did or intended to do from what s/he apparently did, to the extent that this can be discerned in the research report. They confuse the adequacy of a description of something in a report with the appropriateness of something

that occurred in the study itself, as represented in the report.

A case in point is the sample. Existing criteria frequently do not ask reviewers to differentiate between an informationally adequate description of a sample and a sample adequate to support a claim to informational redundancy. In the first instance, the writer has given either enough or not enough information about the sample to evaluate it. In the second instance, the reader makes a judgment that the sample is or is not large enough to support a claim the writer has made. In the first instance, a judgment is made about adequacy of reporting. In the second instance, a judgment is made about the appropriateness of the reported sample itself (that is, its size and configuration) to support the findings. In other words, a judgment of reporting adequacy has to be made before a judgment of procedural or interpretive appropriateness can be made. Before a reader can make a judgment about anything, the writer must have given enough information about it in her/his report. A judgment of procedural or interpretive appropriateness (i.e., is it *good* or *valid*?) presumes a judgment of reporting adequacy (i.e., is it on the page?).

Actual Versus Virtual Presence or Absence

Yet the reader must also have an appreciation for the reporting constraints that may have been placed on the writer, such as page limitations and journal and disciplinary conventions concerning what needs to be explicitly said, what can be implied, and what can be omitted. The absence of something in a report does not mean the absence of that thing in the study itself.

Moreover, readers themselves will vary in their willingness to accept a reporting absence. While reviewing the five studies we had selected for the expert panel, one panel member reported that she was most influenced by the absence of information in a report. As she explained it, if a researcher said nothing about method, then method became highly influential in how she viewed the study. Another reviewer suggested a presence/absence calculus in that the presence of findings "with grab" could favorably offset for her the absence of a well-defined problem or method.

Just as the absence of something in a report does not necessarily mean it was absent in the study itself, so too the presence of something in a report does not necessarily mean it was present in the study itself. A writer may have reported that s/he used phenomenological methods, but the reader — in her or his judgment of what constitutes phenomenology — finds no discernible evidence of the use of those methods in the findings. Instead, the reader finds discernible evidence that the technique used to analyze the data was a form of content analysis. Accordingly, a more appropriate way for the reader to read the study is as a content analysis, not as a phenomenological analysis, despite what the writer reported. The reader is here rewriting the report to conform to her/his reading of it and is arguably, in the process, giving the report a better reading and the writer of the report a reviewing break. Read as a content analysis, the report may be judged a good example of its kind; read as a phenomenological analysis, it may not be.

The description of a procedure may be judged informationally adequate but informationally and/or procedurally inappropriate. A writer may adequately describe the inter-rater reliability coding technique used to validate study findings, but the reader may judge the rendering of the technique itself as inaccurate and/or the actual use of such techniques as inappropriate to the narrative claims made in a study. In addition, a writer may be forced to discuss matters inappropriate to a qualitative study. The best case in point is the frequent discussion of the so-called limitations of qualitative research, where writers may be forced by peer reviewers or

editors to state that their sample was not statistically representative or that their findings are not generalizable. Such statements suggest that a researcher/writer does not understand the purpose of sampling in qualitative research, nor the fact that idiographic and analytic generalizations are outcomes of qualitative research. Yet such statements may not be reflective of any error on the part of the researcher.

In summary

In summary, readers must have a keen grasp of the diversity in qualitative research and of writing styles and constraints. A critical error would be to exclude from consideration studies with valuable findings for reasons unlikely to invalidate the findings. A study presented as a phenomenology that is evidently a qualitative descriptive study may still be a "good" study: that is, a study with credible and useful findings. Another critical error would be to accept what a writer says at face value without looking behind the face. The appraisal of qualitative studies requires discerning readers who know and take account of what their reading preferences are and who are able to distinguish between non-significant representational errors and procedural or interpretive mistakes fatal enough to discount findings. The appraisal of qualitative studies also requires discerning readers able to distinguish between a report that says all of the right things, but which contains no evidence that these things actually took place.

Development, Purpose, and Use of a Guide for Reading Qualitative Studies

The reading guide that follows is the latest version of a tool we developed to assist us in apprehending those features of any one research report most relevant to understanding and ultimately combining its findings with those from a set of reports. We developed it by using the iterative process we described previously involving our own and expert panel members' use and appraisal of successive versions of the guide. We intend the guide to be used with exclusively qualitative and not mixed methods studies, which present distinctive challenges to reading and writing that have been addressed elsewhere (Sandelowski, in press). Although the primary purpose of the guide is to help readers/reviewers read write-ups of qualitative research on health-related topics, it may also be useful to researchers/writers wanting to write up their studies in ways that will appeal to the varied readers in the health sciences.

The purpose of the guide is to make more visible those features of qualitative reports that readers in the health-related practice disciplines are likely to want to see, but which the form of reporting might make it more difficult for them to see. Readers in the health-related practice disciplines typically want information in 13 categories, including research problem, research purpose(s)/question(s), literature review, orientation to the target phenomenon, method, sampling, sample, data collection, data management, validity, findings, discussion, and ethics. Each of these categories is defined in the reading guide shown here. The guide directs readers/reviewers to look for information in these 13 categories, no matter where they might appear in the report. A 14th category — form — asks the reader to consider the general style of the report and, especially, the shape of the findings.

Sometimes bits of information that writers insist are "there" in a report are not seen by the reviewer because they are located in places of the report where reviewers are not looking for them. For example, we have noticed in our review of studies on women with HIV/AIDS that information concerning the ethical and credible conduct of a study was often nested in information provided about the sample, data collection, and data analysis, and/or in the findings and discussion sections. Typically there were no defined sections of the report devoted to the

topics of ethics and validity. Accordingly, the guide helps readers identify what they want to find without rigidly linking that information to any one place in a research report, which might cause a study to be inadequately appraised.

Yet it offers a systematic way to dissect and organize and, therefore, to obtain all of the information available from a report that might subsequently be useful for one or more of the varied purposes readers will have for reading that report, which may be a systematic and comprehensive review of the findings and/or methodologic approaches of a set of studies for a state of the science paper or for a research proposal, or to conduct a metasynthesis. Or, the guide can be used in a more focused way to target key features of qualitative studies in a domain, such as the kinds of participants who have been included or the kinds of recruitment strategies used.

Although the guide is arranged to reflect the topics and order both apparent in most qualitative research reports in the health sciences and expected by most readers of these reports, we intend it to be used dynamically to reflect the purpose of the reading and the nature of the report itself. As the sections of the guide artificially separate and arrange what are integral elements of a whole, some reviewers may want to begin their reading of a report with the findings and sample, some will want to read in the given order of the report, and some will want to read in the given order of the guide. Some categories of information cannot be fully understood until all of the report is read, while other categories may be evident in more self-contained sections of a report. Moreover, any one statement from a research report can be placed in more than one category as it may carry information applicable in more than one category. A statement about how a researcher coded data may be relevant to both the *data management* and *validity* categories.

The purpose for reading a report will determine how the guide is used. Reviews will likely be more detailed when the purpose is to determine whether a study meets criteria for inclusion in a qualitative metasynthesis, but may be less detailed when the purpose is to survey methodological approaches in a field of study. For very detailed work, we recommend using a hard copy of the guide, along with a scanned copy of the research report, from which words, phrases, and/or paragraphs can be copied directly into the template shown in Figure 4.

The guide is thus a reconstruction (of a report) that is itself a reconstruction (of a study); it asks readers to re-shape a report to conform to its logic. By this reshaping effort, the guide makes visible how differently common features of a report contribute to a unique whole. As one panel member suggested, the research report is a "gestalt" with the different components comprising that gestalt variously operating in the foreground or background. Two other panel members suggested that the very same features of a report can enhance or detract from the value assigned to the study it represents. As we learned from our expert panel and from using the guide ourselves, there are at least two effects of this re-shaping effort. One effect is for the reviewer to feel as if s/he were taking a research report apart and rearranging its parts so that it no longer resembles what it was (as Picasso did in many of his paintings of human beings and animals). A contrasting effect is for the reviewer to feel as if s/he were taking a work not immediately recognizable as a research report and rearranging its parts so that it more clearly resembles a scientific research report. No matter which of these or other effects the guide will have on readers, it will compel them to see a report in ways they had not before.

The guide asks readers/reviewers to consider also the presence and, even more importantly, the relevance of specified appraisal parameters. A reader may judge that a category of information

has been addressed — and/or addressed well or badly — but decide that no matter whether or how it was addressed, it did not matter anyway to the overall value of the report. Accordingly, the guide helps readers see what is there, where it is, and what is not there in a report. One panel member noticed how often information about sampling and analysis was missing from reports. And, the guide helps readers better to understand themselves as readers. The guide makes visible what a reader's inclinations are and whether and how they figure into their reviews.

The expert panel members demonstrated this point, in addition to the futility of efforts to create a quantitatively reliable tool to appraise qualitative studies. We asked the expert panel members to rank how important (1 as most important and 14 as least important) each of the 14 categories of information were to them in reviewing each of five qualitative studies on women with HIV/AIDS. The panel members uniformly agreed that it was virtually impossible reliably to rank all 14 categories, one member specifically objecting to any effort to quantify what for her is ultimately a qualitative assessment. As the panel members observed, it was difficult to disaggregate parts that were not only integrally connected to each other, but also connected to each other in different ways in each of the five studies. However, they did find it easier and more acceptable to select the three categories that influenced them the most, and the three categories that influenced them the least, in appraising each study. Their rankings are summarized in Figures 1-3.

As shown in these Figures, reviewers showed relatively low inter-rater or intra-rater consistency in the categories of information they selected as most influential and least influential across and within ranking positions and across studies. But they did show some reading/ reviewing tendencies or preferences. As shown in Figure 1, Reviewer A tended to select *method*, *data management*, and *validity* as the most influential, and *discussion* as the least influential categories. Reviewer B showed few preferences in her most-influential rankings, but showed a preference for *ethics* as least influential. Reviewer C demonstrated a preference for *findings* as the category most influential to her, and a preference for *discussion* and *ethics* as the categories least influential to her. Reviewer C also demonstrated that a category — in her case, *orientation to target phenomenon* — could be ranked by the same reviewer as most influential in one study and least influential in another. Reviewer D showed a preference for *findings* as most influential, and for *ethics* and *data management* as least influential. Finally, Reviewer E showed a preference for *findings* and *purpose* as the most influential, and *discussions* and *ethics* as the least influential categories. Figure 3 shows an overall preference for ranking *findings* as most influential and *discussion* and *ethics* as least influential. Figure 2 suggests that reports themselves may elicit preferences. For example, we surmise that the reason Study #4 elicited *target* as among the most influential categories was that it was an anthropological study with a prominent and highly distinctive theoretical orientation to its target phenomenon that was the single most important determinant of the findings and the single most characteristic feature of the report as a whole.

In summary, we see the guide as itself a literary technology designed to enhance the appreciation and thereby promote a more informed appraisal of the literary technology we know as the qualitative research report. We offer it as a "conceptual and presentation device" (Maxwell, 1996, p. 8), not as a set of rules to be slavishly followed. We agree with Maxwell (1996, p. ix) that "a guide. . . is best when those guiding you are opinionated." This guide certainly derives from our opinions, but these opinions are grounded in our current research and years of experience in reading and writing qualitative research. We fully anticipate and invite the readers/reviewers who use this guide to exercise their opinions and, in the process, to

enhance the quality and utility of the guide. Any guide is always a work-in-progress.

Notes

1. This study, entitled "analytic techniques for qualitative metasynthesis," is supported by grant # R01 NR04907 from the National Institute for Nursing Research.
2. They are Cheryl Tatano Beck, Louise Jensen, Margaret Kearney, George Noblit, Gail Powell-Cope, and Sally Thorne.
3. One panel member could not attend because of inclement weather.
4. Another panel member is currently on sabbatical and did not participate in this exercise.
5. Material in this portion of the text is also included in a different and expanded form in a chapter previously prepared for an anthology that will be published as Sandelowski, In press, cited in the reference list.

References

- Amann, K., & Knorr-Cetina, K. (1988). The fixation of visual evidence. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 85-121). New York: Kluwer Academic.
- Ayres, L., & Poirier, S. (1996). Virtual text and the growth of meaning in qualitative analysis. *Research in Nursing & Health, 19*, 163-169.
- Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison, WI: University of Wisconsin Press.
- Beach, R. (1993). *A teacher's introduction to reader-response theories*. Urbana, IL: National Council of Teachers of English.
- Bochner, A. P. (2000). Criteria against ourselves. *Qualitative Inquiry, 6*, 266-272.
- Clifford, J., & Marcus, G. E. (Eds.). (1986). *Writing culture: The poetics and politics of ethnography*. Berkeley, CA: University of California Press.
- Clark, G. (1990). *Dialogue, dialectic, and conversation: A social perspective on the function of writing*. Carbondale, IL: Southern Illinois University Press.
- Devers, K. J. (1999). How will we know "good" qualitative research when we see it? Beginning the dialogue in health services research. *HSR: Health Services Research, 34*, 1153-1188.
- Dillon, A., & Vaughan, M. (1997). "It's the journey and the destination": Shape and the emergent property of genre in evaluating digital documents. *New Review of Multimedia and Hypermedia, 3*, 91-106.
- Eisner, E. (1985). Aesthetic modes of knowing. In E. Eisner (Ed.), *Learning and teaching the ways of knowing: Eighty-fourth yearbook of the National Society for the study of Education, Part II* (pp. 23-36). Chicago: National Society of the Study of Education.

- Eisner, E. (1991). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. New York: Macmillan.
- Emden, C. & Sandelowski, M. (1998). The good, the bad, and the relative, part 1: Conceptions of goodness in qualitative research. *International Journal of Nursing Practice*, 4, 206-212.
- Emden, C. & Sandelowski, M. (1999). The good, the bad, and the relative, part 2: Goodness and the criterion problem in qualitative research. *International Journal of Nursing Practice*, 5, 2-7.
- Engel, J. D., & Kuzel, A. J. (1992). On the idea of what constitutes good qualitative inquiry. *Qualitative Health Research*, 2, 504-510.
- Fish, S. (1980). *Is there a text in this class? The authority of interpretive communities*. Cambridge, MA: Harvard University Press.
- Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, 3, 430-441.
- Garratt, D., & Hodkinson, P. (1998). Can there be criteria for selecting research criteria? A hermeneutical analysis of an inescapable dilemma. *Qualitative Inquiry*, 4, 515-539.
- Geertz, C. (1988). *Works and lives: The anthropologist as author*. Stanford, CA: Stanford University Press.
- Gephart, R. P. (1986). Deconstructing the defense for quantification in social science: A content analysis of journal articles on the parametric strategy. *Qualitative Sociology*, 9, 126-144.
- Gephart, R. P. (1988). *Ethnostatistics: Qualitative foundations for quantitative research*. Beverly Hills, CA: Sage.
- Golden-Biddle, K., & Locke, K. (1993). Appealing work: An investigation of how ethnographic texts convince. *Organization Science*, 4, 595-616.
- Gusfield, J. (1976). The literary rhetoric of science: Comedy and pathos in drinking driver research. *American Sociological Review*, 41, 16-34.
- Heron, J. (1996). Quality as primacy of the practical. *Qualitative Inquiry*, 2, 41-56.
- Hunter, A. (Ed.). (1990). *The rhetoric of social research: Understood and believed*. New Brunswick, NJ: Rutgers University Press.
- John, I. D. (1992). Statistics as rhetoric in psychology. *Australian Psychologist*, 27, 144-149.
- Kearney, M. H. (2001). Levels and applications of qualitative research evidence. *Research in Nursing & Health*, 24, 145-153.

- Kvale, S. (1995). The social construction of validity. *Qualitative Inquiry*, 1, 19-40.
- Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.
- Law, J., & Whittaker, J. (1988). On the art of representation: Notes on the politics of visualization. In G. Fyfe & J. Law (Eds.), *Picturing power: Visual depiction and social relations* (pp. 160-183). London: Routledge.
- Lincoln, Y. S., & Reason, P. (Eds.). (1996). Quality in human inquiry (Entire issue). *Qualitative Inquiry*, 2 (1).
- Lye, J. (1996a). *Reader-response: Various positions*. Available on line at www.brocku.ca/english/courses/4F70/rr.html
- Lye, J. (1996b). *Some factors affecting/effecting the reading of texts*. Available on line at www.brocku.ca/english/courses/4F70/factors.html
- Lynch, M., & Woolgar, S. (Eds.). (1988). *Representation in scientific practice*. New York: Kluwer Academic.
- Lynch, M., & Edgerton, S. Y. (1988). Aesthetics and digital image processing: Representational craft in contemporary astronomy. In G. Fyfe & J. Law (Eds.), *Picturing power: Visual depiction and social relations* (pp. 184-220). London: Routledge.
- Manguel, A. (1996). *A history of reading*. Toronto: Alfred A. Knopf.
- Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, 62, 279-300.
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.
- McGill, L. T. (1990). Doing science by the numbers: The role of tables and other representational conventions in scientific journal articles. In A. Hunter (Ed.), *The rhetoric of social research: Understood and believed* (pp. 129-141). New Brunswick, NJ: Rutgers University Press.
- Norris, J. R. (1997). Meaning through form: Alternative modes of knowledge representation. In J. M. Morse (Ed.), *Completing a qualitative project: Details and dialogue* (pp. 87-115). Thousand Oaks, CA: Sage.
- Popay, J., Rogers, A., & Williams, G. (1998). Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research*, 8, 341-351.
- Richardson, L. (2000a). Evaluating ethnography. *Qualitative Inquiry*, 6, 253-255.
- Richardson, L. (2000b). Writing: A method of inquiry. In N. K. Denzin & Y. S. Lincoln

(Eds.), *Handbook of qualitative research*, (2nd ed.) (pp. 923-948). Thousand Oaks, CA: Sage.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work*. Carbondale, IL: Southern Illinois University Press.

Sandelowski, M. (In press). Tables or tableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Schwalbe, M. (1995). The responsibilities of sociological poets. *Qualitative Sociology*, 18, 393-413.

Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.) (pp. 189-213). Thousand Oaks, CA: Sage.

Seale, C. (1999). *The quality of qualitative research*. London: Sage.

Shapin, S. (1984). Pump and circumstance: Robert Boyle's literary technology. *Social Studies of Science*, 14, 481-520.

Sparkes, A. C. (2001). Myth 94: Qualitative health researchers will agree about validity. *Qualitative Health Research*, 11, 538-552.

Stanley, L. D. (1999). Transforming AIDS: The moral management of stigmatized identity. *Anthropology & Medicine*, 6, 103-120.

Tierney, W. G. (1995). (Re)Presentation and voice. *Qualitative Inquiry*, 1, 379-390.

Van Maanen, J. (1988). *Tales of the field: On writing ethnography*. Chicago: University of Chicago Press.

Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research*, 11, 522-537.

Figure 1. Expert panel member rating of most and least influential information categories

Reviewer A						
Studies	Study 1	Study 2	Study 3	Study 4	Study 5	
Most influence						Percent agreement
1	Method	D manage	Method	Method	Method	80%

2	Target	D collect	D manage	D manage	D manage	60%
3	D manage	Sampling	Validity	Validity	Validity	60%
Least influence						
12	Literature	Ethics	Sampling	Form	Discussion	0
13	Form	Discussion	Sample	Ethics	Literature	0
14	Discussion	Literature	Discussion	Discussion	Form	60%

Reviewer B						
Studies	Study 1	Study 2	Study 3	Study 4	Study 5	
Most influence						Percent agreement
1	D collect	D manage	Problem	Findings	Findings	40%
2	D manage	Findings	Purpose	Form	Validity	0
3	Validity	Literature	Literature	Discussion	D collection	40%
Least influence						
12	Discussion	Form	D manage	D manage	Sampling	40%
13	Form	Discussion	Form	Validity	Sample	40%
14	Ethics	Ethics	Ethics	Ethics	Ethics	100%

Reviewer C						
Studies	Study 1	Study 2	Study 3	Study 4	Study 5	
Most influence						Percent agreement
1	Method	Findings	Findings	Findings	Findings	80%
2	Purpose	Purpose	Purpose	Method	Method	Split 60/40%

3	Target	Target	Problem	Target	D collect	60%
Least influence						
12	Form	Discussion	Target	Ethics	Discussion	40%
13	Ethics	Ethics	Form	D collect	Form	Split 40/40%
14	Discussion	Form	Discussion	D manage	Ethics	40%

Reviewer D						
Studies	Study 1	Study 2	Study 3	Study 4	Study 5	
Most influence						Percent agreement
1	Findings	Findings	Findings	Findings	Form	80%
2	Target	Problem	Sample	Discussion	Discussion	40%
3	Literature	Purpose	Problem	Target	Target	40%
Least influence						
12	Sampling	Method	D collection	D manage	Purpose	0
13	Ethics	Validity	D manage	Validity	Problem	40%
14	Purpose	D manage	Ethics	Ethics	Ethics	60%

Reviewer E						
Studies	Study 1	Study 2	Study 3	Study 4	Study 5	
Most influence						Percent agreement
1	Findings	Findings	Findings	Purpose	D manage	60%
2	Sample	Target	Sample	Findings	Findings	Split 40/40%

3	Purpose	Method	Purpose	Target	Purpose	60%
Least influence						
12	D manage	D collect	Discussion	D collect	Sampling	40%
13	Ethics	Sampling	Validity	Discussion	Ethics	40%
14	Literature	Discussion	D manage	Ethics	Literature	40%

D = data

Figure 2. Most and least influential rankings of 14 categories in five studies (n=5)

Studies		1	2	3	4	5
Ranks	Reviewer					
1	A	Method	D manage	Method	Method	Method
	B	D collect	D manage	Problem	Findings	Findings
	C	Method	Findings	Findings	Findings	Findings
	D	Findings	Findings	Findings	Findings	Form
	E	Findings	Findings	Findings	Purpose	D manage
	percent agree	split 40% / 40%	split 60% / 40%	60%	60%	40%
2	A	Target	D collection	D manage	D manage	D manage
	B	D manage	Findings	Purpose	Form	Validity
	C	Purpose	Purpose	Purpose	Method	Method
	D	Target	Problem	Sample	Discussion	Discussion
	E	Sample	Target	Sample	Findings	Findings
	percent	40%	0	split	0	0

	agree			60% / 40%		
3	A	D manage	Sampling	Validity	Validity	Validity
	B	Validity	Literature	Literature	Discussion	D collect
	C	Target	Target	Problem	Target	D collect
	D	Literature	Purpose	Problem	Target	Target
	E	Purpose	Method	Purpose	Target	Purpose
	percent agree	0	0	40%	60%	40%
12	A	Literature	Ethics	Sampling	Form	Discussion
	B	Discussion	Form	D manage	D manage	Sampling
	C	Form	Discussion	Target	Ethics	Discussion
	D	Sampling	Method	D collect	D manage	Purpose
	E	Purpose	D collect	Discussion	D collect	Sampling
	percent agree	0	0	0	0	split 40% / 40%
13	A	Form	Discussion	Sample	Ethics	Literature
	B	Form	Discussion	Form	Validity	Sample
	C	Ethics	Ethics	Form	D collect	Form
	D	Ethics	Validity	D manage	Validity	Problem
	E	Ethics	Sampling	Validity	Discussion	Ethics
	percent agree	split 60% / 40%	40%	40%	40%	0
14	A	Discussion	Literature	Discussion	Discussion	Form
	B	Ethics	Ethics	Ethics	Ethics	Ethics
	C	Discussion	Form	Discussion	D manage	Ethics
	D	Purpose	D manage	Ethics	Ethics	Ethics

	E	Literature	Discussion	D manage	Ethics	Literature
	percent agree	40%	0	split 40% / 40%	60%	60%

D = data

Figure 3: Frequency of ratings of 14 information categories in five studies (n=5)

Rankings	1	2	3	12	13	14
<u>Categories</u>						
Problem	1	1	2	0	1	0
Purpose	1	4	3	1	0	1
Literature	0	0	3	1	1	3
Target	0	3	6	1	0	0
Method	5	2	1	1	0	0
Sampling	0	0	1	4	1	0
Sample	1	2	0	0	2	0
D collect	1	1	2	3	1	0
D manage	3	5	1	4	1	3
Validity	0	1	4	0	4	0
Findings	13	3	0	0	0	0
Discussion	0	2	1	5	3	6
Ethics	0	0	0	2	6	10
Form	1	1	0	3	5	2

A guide for Reading Qualitative Studies

Face page
Create an inventory of the demographic features of and reading context for the article. This will help you identify the manifest features and historical context of a report, and the purpose for which you are reading this report at this time. For example, is it to prepare a research proposal, to chart the state of the science in a field, to identify methodological approaches

used in a field, or to conduct a metasynthesis study? Dating the study will also help you evaluate the clinical relevance of findings.

Demographic Features

Complete citation

Author affiliations, including discipline and institution

Funding source

Acknowledgment

Period of data collection

Dates of submission and acceptance of work

Publication type (e.g. authored/edited book, journal, dissertation, thesis, conference proceeding)

Mode of retrieval (e.g. computer database, citation list, personal communication)

Key words (in article and by reviewer)

Abstract

Reading Context

Date of reading

Purpose of reading

Reader

Reader affiliations

Problem

Look for information concerning what is wrong, or missing, or needed that requires fixing, finding, or satisfying. The research problem is usually a clinical problem in the practice disciplines, and a theoretical or disciplinary problem in the social science disciplines. An example of a clinical problem is:

Many women with HIV wait too long to obtain treatment. Delays in obtaining HIV-related treatment have been linked to shorter survival for women after diagnosis. These delays must be stopped, but we do not know enough about why they occur.

An example of a theoretical problem is:

Stigma has generally been conceived as a negative event. But there are circumstances in which stigma has positive outcomes. Theories of stigma should be expanded to include these positive outcomes.

Generally appearing early in the experimental style research report, problem statements set the stage for the study that was conducted and typically establish the significance of and/or reason for the research purpose. Problems may be explicitly stated or they may be implied in the research purpose and/or the literature review.

Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. There is a discernible problem that led to the study.</p> <p>2. The problem is accurately depicted.</p> <p>3. The problem is related to the research purpose and/or the literature review.</p> <p>4. The problem establishes the significance of the research purpose, or why the researcher wanted to conduct the study, beyond simply stating that "no one has studied this (qualitatively) before."</p>		

Purpose(s)/question(s)		
<p>Look for statements concerning one or more goals, objectives, or aims of the study, and/or a list of one or more questions the study findings will answer. Research purposes may be explicitly stated, or they may be apparent in statements such as "I intend/hope to show..." or "I will argue/suggest..." In more ethnographically styled reports, implied statements of purpose may be found in the foreshadowing or summarizing of the findings early in the report that will be described in more detail later in the report. For example:</p> <p>In this article, I will show how white and middle-class women with HIV/AIDS morally manage a stigmatized identity. I suggest that the reclamation of a coherent and safe moral identity is an integral part of how they cope with seropositivity and manage stigma...(Stanley, 1999, p. 104)</p>		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. There is a discernible set of research purposes and/or questions.</p> <p>2. Research purposes or questions are linked to the problem and/or to the review of the literature.</p> <p>3. Research questions are amenable to qualitative study.</p>		

Literature review
<p>Look for information concerning what is believed, known, and not known about a problem. Sometimes the literature reviewed is combined with information about the problem, while other times, it is set off in a separate section and labeled as a literature review or with the actual topics contained in the review. In addition, reviews of literature may relate to the findings researchers will feature in the report, as opposed to the problem that originally led to</p>

the study.

Reviews of literature may show one or a combination of the following logics:

1. A deficit/gap logic where writers emphasize what is not known about a problem and point to a purpose that will offset this knowledge deficit
2. An error logic where writers emphasize what is mistaken about what is "known" and point to a purpose that will correct this error
3. A contradiction logic where writers emphasize the inconsistencies in knowledge and point to a purpose that will help to resolve this contradiction and/or
4. A synthesis logic where writers emphasize the common areas in two or more seemingly disparate bodies of empirical, theoretical, or other literature and point to a purpose that will illuminate this overlap.

Appraisal parameters	Presence Yes/No	Relevance Yes/No
<ol style="list-style-type: none"> 1. Key studies and other relevant literatures addressing the research problem are included. 2. The review is related to the research problem. 3. The review clarifies whether it reflects what researchers know and believed going into the field of study — before any data were collected — or came to know and believe while in the field coming out of it, after data analysis began or was completed. 4. The review shows a critical attitude, as opposed to simply and/or indiscriminately summarizing studies. 5. The review shows a discernible logic that points toward the research purpose. 		

Mindset toward target phenomenon

Look for indications of the perspectives, assumptions, conceptual/theoretical frameworks, philosophies and/or other frames of reference, mindsets, or "theoretical sensitivities" guiding or informing researchers concerning the target phenomenon or subject matter of a study: i.e., the people, events, or things to be studied. For example, Goffman's theory of stigma is used to frame a study of HIV+ women's social interactions. Such frames of reference may be explicitly stated, as in the Goffman example. Or, they may be implied in the language used to

describe the target phenomenon, and/or in the literature reviewed about it, and/or in the problem identified concerning it, or the questions asked about it, as when HIV+ women's responses to infection are discussed in terms of "self-care" or "coping," and studies in self-care or coping are reviewed. Such frames of reference may be clearly distinguishable from the methodological location of a study, or overlap with it. For example, feminism may be presented as the framework for the study of women's responses to HIV diagnosis in particular, and/or as the framework for any study of women and/or for inquiry in general. A frame of reference may have influenced a study from its conception through the interpretation of findings. In contrast, a frame of reference may not have entered the study until after some or all of the data were collected and analyzed. For example, Goffman's ideas about stigma may have been the a priori or sensitizing framework for a study of women with HIV. That is, these women are seen from the beginning through to the end of the study as living with and responding to a culturally stigmatizing condition. In contrast, Goffman's ideas might have entered a study only after researchers had begun to analyze their data and recognized that women's responses fit and/or were illuminated by these ideas.

Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. There is an explicitly stated or implied frame of reference.</p> <p>2. If explicitly stated, the frame of reference is accurately rendered.</p> <p>3. Whether stated or implied, the frame of reference fits the target phenomenon. That is, it is not forced onto the target phenomenon, as when a theory emphasizing other people's knowledge of a stigmatizing condition as critical to the way the person having the condition experiences it is used to frame the experiences of a group of HIV+ women who never disclosed their condition to others.</p> <p>4. If explicitly stated as the guiding frame of reference for a study, it played a discernible role in the way the study was conducted and/or the way the findings were treated. This is in contrast to a frame of reference that is evidently operating in a study, but which is not demonstrably recognized by the researcher as when HIV+ women are consistently referred to as being "in denial," but denial as a concept is never discussed nor recognized for its interpretive impact. Or, the researcher does not recognize that s/he is viewing self-care as activities health care providers view as positive and not as encompassing such activities as smoking and drug abuse, which can also be construed as self-care.</p> <p>5. The presentation of the mindset for the study clarifies whether it influenced researchers going into the field of study — before any data were collected — or while in the field or coming out of it, after data analysis began or was completed.</p>		

Method		
<p>Look for indications of the perspectives, assumptions, philosophies and/or other frames of reference guiding or informing researchers concerning the conduct of a study. For example, grounded theory is presented as the method and as deriving from tenets of symbolic interactionism and pragmatism. Semiotics is presented as the analytic frame of reference for the study of a document or artifact. Such frames of reference may be explicitly stated, or implied in the method language and/or citations used. For example, no method may ever be named per se, but phrases such as "lived experience," suggesting phenomenology, and "theoretical sampling," suggesting grounded theory, are used; and/or there are citations to Van Manen's work on phenomenology or Strauss & Corbin's work on grounded theory. The method location of a study may be clearly distinguishable from the conceptualization of the target phenomenon of a study, or it may overlap with it. For example, social constructionism may be presented as the framework for any study of women and/or for inquiry in general, and for a study of women's responses to HIV diagnosis in particular.</p>		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<ol style="list-style-type: none"> 1. There is a stated or implied method. 2. The method fits the research purpose. 3. The method is accurately rendered. 4. The uses of method-linked techniques for other than method-linked purposes are explained as when theoretical sampling is used in a qualitative descriptive study, or phenomenological techniques are used to create items for an instruments. 		

Sampling strategy & techniques		
<p>Look for information about researchers' sampling intentions going into a study and evolving sampling decisions in the course of the study, including planned recruitment sites.</p>		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<ol style="list-style-type: none"> 1. The sampling plan fits the purpose and method. 2. The sampling plan is purposeful. 3. The sampling plan described is accurately rendered, as opposed to being inaccurately rendered or misrepresented as when maximum variation sampling is presented as having equal 		

numbers of men and women, or percents of African Americans or Hispanic Americans equal to their presence in the population.		
4. Sites of recruitment fit the purpose and sampling strategy.		

Sample		
<p>Look for a description of the people (including the configuration of focus groups), places, events, documents, and/or artifacts comprising the actual sources of information for the study, and the actual sites from which people were recruited. Because ethnographic studies are typically site/place-bound, the site is actually a component of the sample. Site—as sample—is contrasted with site of data collection. That is, a study may involve one organization (site as sample), and interviews may be conducted in conference rooms on site (site of data collection).</p>		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. Sample size and configuration fit the purpose and sampling strategy.</p> <p>2. Sample size and configuration can support claims to informational redundancy, or theoretical or scene saturation.</p> <p>3. Sample size and configuration can support claims to intensive, comprehensive, or holistic studies in particular.</p> <p>4. Sample size and configuration can support the findings.</p> <p>5. The sample is presented in a case-oriented way, as opposed to a variable-oriented was as when, in the report of a study of mothering in 20 HIV+ women, menas and ranges are given for the numbers of pregnancies and children who were also HIV+, but the unique combination of these variables in each mother-child dyad is not shown or addressed anywhere in the report.</p> <p>6. Features of the sample critical to the understading of findings are described, as opposed to not describes as when, in a study of HIV+ women's reproductive decision-making, no information is offered on women's use of contraceptives, obstetrics histories, no on severity of disease.</p> <p>7. Sites of recruitment fit the evolving needs of the study.</p>		

Data collection techniques & sources		
<p>Look for a description of the techniques and procedures used to obtain information for a study in one or more of the following categories: interviews (including focus groups), observations, documents, and artifacts. Look for descriptions of the purpose and place of interviews or observations, the type of, orientation to, and/or manner of conducting interviewing, observation, document review, or artifact study, and of the timing and sequencing of data collection. Look also for information about alterations in techniques and procedures made in the course of the study.</p>		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. Data collection techniques and sources fit the purpose and mindsets of the study, as opposed to not fitting them as when the purpose of a study is to ascertain structural barriers to health care utilization, but the only source of data is women's perceptions of their health care providers. Or, researchers conflate the longitudinal with the validation purpose for conducting more than one interview with the same participants or more than one observation of the same event.</p> <p>2. Specific data collection techniques are tailored to the reported study, as opposed to the presentation of textbook or rote descriptions of data collection with no application shown to the study reported.</p> <p>3. Data collection techniques are accurately rendered, as opposed to inaccurately rendered as when the observation process that occurs during interviews and focus groups is presented as participant observation.</p> <p>4. The sources of data presented are demonstrably the basis of the findings, as opposed to not being their basis as when document study is presented as a data collection strategy, but there is no evidence of its use.</p> <p>5. Data collection techniques are correctly used, as opposed to misused as when focus groups are conducted by asking each participant in turn to answer the same question, instead of posing a question to the group to stimulate group interaction.</p> <p>6. The sequence and timing of data collection strategies vis-à-vis each other fit the purpose and mindsets of the study.</p> <p>7. Sites are conducive to data collection.</p>		

8. Alterations in techniques fit the evolving needs of the study.		
9. The time period for data collection is explicitly stated.		

Data management techniques		
<p>Look for a description of techniques used to 1) create data; 2) create an audit trail of data; 3) prepare data for analysis; 4) catalogue, file, or organize data sets; and 5) break up, (dis)play (with), or reconfigure data. Included here is information on whether and how transcripts of interviews and field notes were prepared, whether and which computerized text management systems were used, the specific analytic approaches employed (e.g., content, constant comparison, narrative, discourse, or other analysis), and whether and how data matrices and other visual displays of data were used. Information about these techniques may be explicitly stated, or shown or implied in the findings.</p>		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. Data management techniques fit the purposes and data.</p> <p>2. Specific data management techniques are tailored to the reported study, as opposed to textbook or rote descriptions of data management being offered, with no application shown to the study reported.</p> <p>3. Data management techniques are accurately rendered.</p> <p>4. Data management techniques are correctly used.</p> <p>5. There is a clear plan for analytically linking interview, observation, document, and/or artifact data sets.</p>		

Findings		
<p>Look for what researchers "found" from the data they collected, or the results of their interpretation of these data. Findings are to be distinguished from data, or the case descriptions, field notes, or quotes that support an interpretation. Findings will show varying levels of complexity, from a basic descriptive summary to a highly interpreted conceptual rendering.</p>		
Appraisal parameters	Presence	Relevance

	Yes/No	Yes/No
<p>1. There is a discernable set of results distinguishable from the data researchers collected, as opposed to indistinguishable as when the researcher presents several case histories but offers no interpretation of them. This is an example of descriptive excess or heaped description, as opposed to thick description.</p> <p>2. The results of the study are distinguishable from the researcher's discussion of the results or from the results of other studies to which the researcher refers.</p> <p>3. Interpretations of data are demonstrably plausible and/or sufficiently substantiated with data, as opposed to implausible as when a mother is quoted as hitting her child and this quote is used to illustrate the "joys of motherhood".</p> <p>4. Data are sufficiently analyzed and interpreted.</p> <p>5. Findings address the research purpose, as opposed to not addressing them as when the stated purpose of a study was to describe structural barriers to health care utilization, but the findings focus on women's perceptions of their health care.</p> <p>6. Variations in sample and/or data are addressed.</p> <p>7. Analysis is largely case-oriented, or oriented to the study of particulars, as opposed to variable-oriented or quantitatively-informed.</p> <p>8. Concepts or ideas are well-developed and linked to each other.</p> <p>9. Concepts are used precisely, as opposed to imprecisely as when sources of social support are persistently conflated with perceptions of others as supportive.</p> <p>10. Analysis of data fits the data, as opposed to not fitting as when focus group data are analyzed at the individual level and the analysis takes no account of group interaction.</p> <p>11. The results offer new information about, insight into, or formulation of the target phenomenon.</p> <p>12. The findings are relevant for contemporary use, as opposed to being irrelevant as when data from HIV+ women were collected pre-HAART and when AIDS was considered a fatal as opposed to chronic disease.</p>		

Discussion		
Summary of and conclusions about the findings of the study, and a discussion of their clinical, theoretical, policy, disciplinary, or other significance.		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. Discussion of findings is based on the study findings previously described, as opposed to being contrary to the findings, or to introducing findings not previously described.</p> <p>2. The study findings are linked to findings in other studies, or to other relevant literatures either previously discussed or newly introduced.</p> <p>3. The clinical, policy, theoretical, disciplinary, and/or other significance of the findings is thoughtfully considered, as opposed to indiscriminately considered as when changes in practice are recommended that merely propose actions opposite to the findings (e.g., providers are found to be insensitive so the implication is that they must be educated to become sensitive), or when repeating a study with other populations and/or in other settings is recommended with no rationale.</p>		

Validity		
Look for discussions of techniques specifically intended to ensure that the study is scientifically and/or ethnographically valid or "good". Included is information about the strengths and limitations of a study, of specific topics such as reflexivity, reliability, rigor, credibility, and plausibility, and of specific procedures, such as member validation and peer review. Information about validity may be explicitly stated, or implied in discussions of sampling, the sample, data collection and analysis, and in the presentation of the findings. Researchers may emphasize, although not identify as such, different kinds of "validities" in their study: e.g., descriptive, interpretive, theoretical, and pragmatic validity.		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. Researchers show an awareness of their influence on the study and its participants.</p> <p>2. The distinctive limitations of the study are summarized: e.g., theoretical sampling could not be fully conducted in a grounded theory study. This is in contrast to summarizing and/or</p>		

<p>apologizing for the so-called limitations of qualitative research.</p> <p>3. Techniques for validation are used that fit the purpose, method, sample, data, and findings, as opposed to using techniques that do not fit as when reliability coding to ascertain consistency in interview data is used in a study emphasizing the revisionist nature of narratives.</p> <p>4. Techniques used are tailored to the reported study, as opposed to presentations of textbook or rote descriptions of validation techniques with no application shown to the study reported.</p> <p>5. Techniques for validation are accurately rendered, as opposed to misrepresented as when descriptive validity is confused with interpretive validity, and triangulation for convergent validity is confused with using different data sources for completeness.</p> <p>6. Techniques for validation are correctly used, as opposed to incorrectly used as when cases are kept in or dropped from consideration because they conform or do not conform to other cases.</p>		
--	--	--

Ethics		
<p>Look for descriptions of any issues and practices relating to the recruitment, retention, and well-being of human participants in a study. Included here is information concerning how participants were approached and enrolled for a study, the informed consent procedures used, the benefits and risks participants were subjected to by virtue of being in the study, the inducements and protections offered them, and the way they responded to participation in the study.</p>		
Appraisal parameters	Presence Yes/No	Relevance Yes/No
<p>1. Benefits and risks distinctive to the study are addressed, as opposed to textbook or rote descriptions of human subjects issues being offered with no description of their particular relevance to the reported study.</p> <p>2. Recruitment and consent techniques were tailored to fit the sensitivity of the subject matter and/or vulnerability of subjects.</p> <p>3. Data collection and management techniques were tailored to fit the sensitivity of the subject matter and/or vulnerability of subjects.</p>		

<p>4. Examples of data provided as evidence to support findings have analytical value and present subjects fairly, as opposed to having only sensational value or presenting subjects unfairly, as when extreme incidents of events are presented when others would do or when quotes are edited that emphasize the lack of education of subjects.</p>		
--	--	--

<p>Form</p>		
<p>Look at the physical format of the entire report. Within the report, look for the literary style and devices used to present the study and its findings. Consider the reporting style (e.g., experimental, ethnographic), the uses of quotes, numbers, cases, and visual displays (e.g., tables, figures, diagrams, photos), the way findings are actually organized, sectioned, and titled, the title of the report, and the use of language, especially metaphor.</p>		
<p>The findings may be presented according to one or more of the following logics:</p>		
<ol style="list-style-type: none"> 1. quantitatively and thematically, by most-to-least prevalent or most-to-least important theme 2. temporally and thematically, with the clock time of the participants as the primary organizing principle and theme as the secondary organizing principle 3. thematically and temporally, with theme as the primary organizing principle and the clock time of the participants as the secondary organizing principle 4. narratively, as a day/week/month/year in the life of participants 5. narratively, as an unfolding drama in the life of participants 6. perspectively (Rashomon effect), by juxtaposing different points of view of participants and/or of researchers 7. polyvocally, by juxtaposing different voices of participants and/or of researchers 8. conceptually, by using sensitizing concepts from extant theory 9. conceptually, by using a grounded theory template for analysis, such as the conditional matrix, typology, or transition format, or set of working hypotheses 10. episodically, emphasizing key moments of an experience 11. archaeologically, with the clock time of researchers as the primary organizing principle to show how the understanding of an event unfolded for them and/or 12. via representative, exemplary, and/or composite cases or vignettes. 		
<p>Appraisal parameters</p>	<p>Presence Yes/No</p>	<p>Relevance Yes/No</p>
<ol style="list-style-type: none"> 1. The overall literary style of the study fits its purpose, method, and findings. 2. Given the reporting style, elements of the research report are placed where readers are likely to find them. 		

<p>3. There is a coherent logic to the presentation of findings.</p> <p>4. Data were organized in ways that do analytic justice to them, as opposed to not doing them justice as when, in rendering of women's experiences with HIV as having physical, psychosocial, and spiritual aspects, highly disparate ideas are dumped into each section because, on the surface, they share physical, psychosocial, and spiritual features.</p> <p>5. Visual displays, quotes, cases, and numbers clarify, summarize, substantiate, or otherwise illuminate findings, as opposed to being at odds with them as when a quote has more ideas in it than featured by the researcher, or a path diagram shows a relationship between variables at odds with the relationship between them depicted in the text.</p> <p>6. The numerical meaning of such terms as "most", "some", "sometimes", and "commonly" is clear.</p> <p>7. The empirical referent for a theme or concept is clear, as opposed to theme being conflated with experience as when a researcher states that five themes emerged from the data instead of stating that women managed their symptoms in one of five ways; or the writer does not clarify whether the themes s/he is discussing are strategies to accomplish a goal, outcomes of having engaged in these strategies, typologies of behavior, or milestones and turning points in a transition.</p> <p>8. Themes or concepts are presented in a comparative and parallel fashion, as opposed to an unparallel manner as when, in a typology, some types are presented as behaviors, while others are presented as character traits, and each type is not compared to every other type.</p> <p>9. Titles of paper and section headers reflect the content in the paper and sections.</p> <p>10. The form fits the audience for whom the report was produced.</p>		
--	--	--

© Sandelowski and Barroso, 2001

Figure 4. Template of reading guide for on-screen work

<p>Face page</p> <p><u>Complete citation:</u></p> <p><u>Author affiliations, including discipline and institution:</u></p> <p><u>Funding source:</u></p>

<p><u>Acknowledgments:</u> <u>Period of data collection:</u> <u>Dates of submission and acceptance of work:</u> <u>Publication type (e.g. authored/edited book, journal, dissertation, thesis, conference proceeding):</u> <u>Mode of retrieval (e.g. computer database, citation list, personal communication):</u> <u>Key words (in article and by reviewer):</u> <u>Abstract:</u></p> <p><u>Date of reading:</u> <u>Purpose of reading:</u> <u>Reader:</u> <u>Reader affiliations:</u></p>

Problem	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4.	Relevance Yes/No (judge as a category)
Purpose	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3.	Relevance Yes/No (judge as a category)
Literature	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4. 5.	Relevance Yes/No (judge as a category)
Target	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4. 5.	Relevance Yes/No (judge as a category)
Method	(Relevant content to be copied or paraphrased)	Presence Yes/No	Relevance Yes/No (judge as a

		1. 2. 3. 4.	category)
Sampling	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4.	Relevance Yes/No (judge as a category)
Sample	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4. 5. 6. 7.	Relevance Yes/No (judge as a category)
Data collection	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4. 5. 6. 7. 8. 9.	Relevance Yes/No (judge as a category)
Data management	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4. 5.	Relevance Yes/No (judge as a category)
Findings	(Relevant content to be copied or paraphrased)	Presence Yes/No 1. 2. 3. 4. 5.	Relevance Yes/No (judge as a category)

		6.	
		7.	
		8.	
		9.	
		10.	
		11.	
		12.	