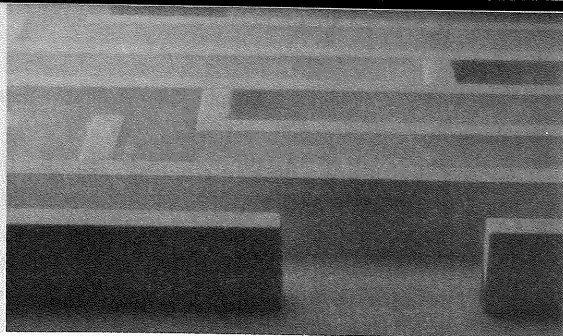


CHAPTER 11



Qualitative quality

Creating a credible, ethical, significant study

Contents

The criteria controversy

Worthy topic

Rich rigor

Sincerity

Credibility

Resonance

Significant contribution

Ethical research practice

Meaningful coherence

In summary

Potential audiences for qualitative work include professors, friends, colleagues, article reviewers, governmental agencies, newspaper editors, policy makers, or lay people. These audiences are drowning in information coming from all directions – newspapers, advertisements, television shows, social networking sites, email accounts, movies, podcasts, blogs, and more. Scholarly reading is just one small slice of this material. Even so, professors review hundreds of student assignments each year, and these compete with exploding numbers of journal articles, policy briefings, memos, strategic reports, and white papers – to say nothing of novels, films, and other media.

With so much information clamoring for attention, a key question is: “How can an inquirer

persuade his or her audiences that the research findings of an inquiry are worth paying attention to?” (Lincoln & Guba, 1985, p. 290). In other words, how do you make your qualitative project attractive, credible, and likely to be taken seriously? Indeed, what ought a qualitative study do? How do we identify high-quality qualitative work?

This chapter addresses such questions and describes how to identify and create quality in qualitative research. First, the chapter reviews common criteria for quality in positivist research and discusses current thinking about the distinguishing characteristics of qualitative goodness. Then it moves into an affirmative discussion of the ways qualitative researchers can strive for valuable, ethical, and inviting work.

The criteria controversy

Devising criteria for scholarly quality is one type of social and humanistic knowledge; therefore such criteria are not “discovered,” but *constructed*. As Guba and Lincoln (2005) advise: “No matter how real, natural, or objective they may seem, criteria are social products created by human beings in the course of evolving a set of practices to which they (and we) subsequently agree to conform” (p. 269). Criteria are human-made filters that necessarily restrict some types of knowledge as they legitimate others; in consequence, criteria are subjective, ever-changing, and sometimes problematic. In order to lay the groundwork for discussing qualitative quality, it is important to understand yardsticks for quality that non-qualitative researchers often use and may mistakenly impose upon qualitative work.

As discussed in Chapter 3, the positivist paradigm still reigns supreme in many scholarly circles. Such an approach assumes a true and empirical reality, complete with knowledge that is “out there,” waiting to be discovered with specific research instruments, which have been validated through replicated use. Good research from this approach connects specific variables with specific outcomes, so as to make predictions and statistical generalizations. Such research aims for **objectivity**, which means that researchers take measures to protect the data and their own analysis of them from being tainted by subjective biases and individual points of view. The notion that research should be “objective” suggests that knowledge-building is best accomplished through measurement devices that are detached from any particular investigator – and that objective scientific procedures will result in the development of facts that can be systematically evaluated.

Another common positivist notion related to quality is **reliability**, which refers to the stability and consistency of a researcher, research tool, or method over time. Reliable studies are those that can be replicated in exactly the same way, no matter who is conducting the study. A reliable instrument always works the same way. For example, a reliable scale measures weight consistently, no matter who is reading it and no matter in what context, who or what is on the scale, and so on. You would not trust a scale that

weighed you at 140 pounds now and at 180 pounds one moment later (unless you happened to pick up a small child or a packed suitcase in between). Reliability is important in the world of scientific measurement – the home of quantitative devices like thermometers and scales.

A third index or criterion of scientific quality is that a study should have **formal generalizability** – a property that refers to the capacity of findings to be transferred from one study to another and to make predictions about how these findings relate to other populations or contexts. For instance, if something happens over and over again (when a baby releases a spoon, it drops to the ground), we predict that it will continue to do so in the future, despite the person's age and context. In order to produce a formal generalization, researchers gather a sufficient amount of data to support a claim in a certain population, they ensure that they have an appropriately randomized sample, and then they statistically calculate how these same findings are true across other populations. By doing all this, they show that the findings of a particular study apply not only to the distinct people envisaged in that study, but also to others. Generalizing always involves a leap of faith, because it predicts future behavior – and also, of course, because it takes knowledge generated in one context and applies it to others, as explained above. However, given a representative (or random) sample containing sufficient data, one can statistically generalize the findings of a smaller group onto another group – in cases like, say, voting patterns.

While these three criteria – objectivity, reliability, and formal generalizability – are very useful in much quantitative research, they do not serve as criteria for most qualitative research – especially the kind that emerges from interpretive, critical, and postmodern approaches. Most qualitative researchers question the very notion of objectivity or consider it to be a **myth** – a powerful story or legend that collectively justifies a certain social practice or institution, but is false or without proof. Seale (1999) argues: "Knowledge is always mediated by preexisting ideas and values, whether this is acknowledged by the researchers or not" (p. 470). Karen Tracy (1995) contends that objectivity is an inappropriate value criterion because it assumes a single world to be known, and therefore it assumes that researchers can actually access the replicability and accuracy of one's observations. I agree; certainly, measures may be taken to reduce one's biases or to account for them, but completely objective and bias-free research is impossible for anyone.

Second, traditional conceptions of reliability have little application to qualitative research, because most qualitative studies are composed of a single analysis, made at a given contextual moment in time. Because socially constructed understandings are always in process and necessarily partial, even if the study were repeated (by the same researcher, in the same manner, in the same context, and with the same participants), the context and participants would have necessarily transformed over time – through aging, learning, or moving on. Hence traditional notions of reliability used in qualitative research are not only mythical, but downright problematic – the "consequence is rather that the study is no good" (Stenbacka, 2001, p. 552).

Formal generalizations – although important for predicting political races, television show ratings, or strains of the flu virus to include in a vaccine – are, similarly, ill suited for qualitative research. This is the case for two reasons. First, most qualitative researchers purposefully trade large or randomized samples for in-depth studies of fewer people or instances – and they often choose to study the unique and the strange rather than the mundane. Therefore the type of sample that would be necessary for reaching formal generalization is rarely desired by qualitative researchers. Second, historically and culturally situated knowledge is ephemeral and always in transformation. Therefore, even if a random

TIPS AND TOOLS 11.1



Eight “big-tent” criteria for excellent qualitative research

Criteria for Quality (end goal)	Various Means, Practices and Methods Through Which to Achieve
Worthy topic	The topic of the research is: • relevant • timely • significant • interesting
Rich rigor	The study uses sufficient, abundant, appropriate and complex • theoretical constructs • data and time in the field • sample(s) • context(s) • data collection and analysis processes
Sincerity	The study is characterized by • self-reflexivity about subjective values, biases, and inclinations of the researcher(s) • transparency about the methods and challenges
Credibility	The research is marked by • thick description, concrete detail, explication of tacit (nontextual) knowledge, and showing rather than telling • triangulation or crystallization • multivocality • member reflections • inter-coder reliability (when collaborating on data-analysis)
Resonance	The research influences, affects, or moves particular readers or a variety of audiences through • aesthetic, evocative representation • naturalistic generalizations • transferable findings
Significant contribution	The research provides a significant contribution • conceptually/theoretically • practically • heuristically • methodologically
Ethical	The research considers • procedural ethics (such as human subjects) • situational and culturally specific ethics • relational ethics
Meaningful coherence	The study • achieves what it purports to be about • uses methods and procedures that fit with its stated goals • meaningfully interconnects literature, research questions/foci, findings, and interpretations with each other

Source: S. J. Tracy, 2010. Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry*, 16, 837–851. Reproduced with permission from Sage

sample could be found, most qualitative researchers would agree that contextualized knowledge, by definition, cannot generalize to other (quite different) scenes in the future.

If positivist criteria are a poor fit for qualitative research, what are qualitative researchers to do? Some qualitative scholars have argued that trying to delineate unvarying research standards is problematic, fruitless, and even silly (Bochner, 2000; Guba & Lincoln, 2005; Schwandt, 1996). I certainly understand that “studies need to be understood and evaluated on their own terms” (Deetz, 2001, p. 38). That said, I believe that criteria can nonetheless be useful in helping us to study, practice, and perfect a method, especially when we are first learning it. Criteria help us to answer the question of whether findings are sufficiently authentic – trustworthy and related to the way others construct their social worlds – and secure – which means that people may act on their implications, for instance to construct social policy or legislation (Guba & Lincoln, 2005).

Hence I developed an expansive “big-tent” approach to criteria for qualitative quality, differentiating the end goals of good research from the mean practices that researchers take to get there (Tracy, 2010). This eight-point conceptualization for obtaining quality in qualitative research serves as a pedagogical tool, promotes dialogue among researchers from various paradigms, and encourages the viability and credibility of qualitative research with a variety of audiences. Tips and Tools 11.1 summarizes the model, and the rest of the chapter discusses how you might design your research so as to attend to these markers of quality.

Worthy topic

The first criterion for qualitative quality is worthiness of the topic. As discussed in earlier chapters, a **worthy topic** can emanate from disciplinary or scholarly theories, relevant or timely social events, or priorities of the particular sample or context of study. Research topics may be worthy because they reveal an aspect of life that has been overlooked, misunderstood, or mistaken, or because they provoke transformation or elicit emotion in the reader.

Particular worth may be found in research that it is counterintuitive, questions taken-for-granted assumptions, or challenges well-accepted ideas. This type of research contrasts with studies that (re)document a phenomenon that is already well established and accepted. Certainly, there is value in strengthening and duplicating studies in order to understand how they may change or remain stable over time. However, worthy studies also point out surprises – issues that shake readers from their commonsense assumptions and practices (Murray, 1971).

Rich rigor

A second marker for quality is **rigor**, which refers to the care and effort taken to ensure that the research is carried out in an appropriate manner. Rigor asks whether researchers have applied due diligence and done their homework. In short, have they put in the time, effort, and thoroughness to practice their craft effectively? The *means* to achieve rigor are multiple and varied. However, rigorous practices include:

- collecting enough data to support significant findings;
- spending enough time in the field to gain trust;
- identifying theoretical goals that are well aligned with your sample or context;
- practicing appropriate procedures in terms of writing fieldnotes, conducting interviews, and analyzing data.

Conducting rigorous research means practicing the discipline and having the motivation to move beyond data and analysis methods that are merely convenient and comfortable. Students completing a course in qualitative methods sometimes want a bit more specificity, so let's consider what qualifies as reasonable parameters for a "rigorous" research assignment in a qualitative methods course.

A scan of recent North American syllabi over the last ten years shows that, depending on the level of the class, students are often asked to spend between 10 and 40 research hours (time to be spent in the field or gathering data from participants) for a course-long project. Qualitative undergraduate projects tend to include about 15 research hours, graduate projects about 25 research hours, master's theses about 80 hours, and dissertations over 100 hours. These numbers reflect what many professors seem to believe is "enough time" to become acquainted with qualitative methods and to create an appropriate study at each student's instructional level.

Despite this "rule of thumb" practice in relation to research hours, there are notable exceptions. If data are new, unique, or rare, not as much time is needed to make a valuable contribution. For instance, two recent articles – one with male executives about work–life balance (Tracy & Rivera, 2010), and another with male professors about being targets of sexual harassment (Scarduzio & Geist-Martin, 2008) – rely on 13 and 4 interviews respectively. Why comparatively so few? Because the data in these studies – despite the limited sample size – are unique. Most studies of work–life balance and sexual harassment are completed with women, and high-ranking men are largely absent from the literature. If these two studies had interviewed women instead of men, they would have likely needed larger interview pools in order to make a significant contribution.

Decisions about how much data to collect also intersect with the "density" of the data. Karen Tracy and I made use of one hundred hours of participant observation data, seven interviews, and a handful of training manuals for the claims we made in our article on emotional labor at 911 (S. Tracy & Tracy, 1998). In contrast, our study of rudeness and face threat during 911 calls relied primarily on *only two* 911 incidents (K. Tracy & Tracy, 1998). The analysis of the data in this second article was *extremely detailed* – taking in the elements of the talk sequence one by one rather than pulling them globally from the emergent themes. When they proceed by this method, some articles may use only a small slice of the data.

Granted, articles using small amounts of data may still emerge from a larger data set. In Scarduzio and Geist-Martin's (2008) study of four sexually harassed male professors, their larger study included 41 interviews. In our 911 research, the two "critical incident" rude calls emerged from an analysis of seven interviews, 100 hours of participant observation, and 650 audio-taped archived calls. Clearly, there is no magic number for the amount of time in the field. The most salient issue to consider is whether the data collected will substantiate meaningful or significant claims.

In terms of rigorous data *analysis*, the reader deserves an explanation about the process by which the raw data were transformed and organized into the research report. As discussed in Chapters 9 and 10, methodology sections should detail the systematic process of sorting, choosing, and organizing the data, whether these operations were accomplished through dialogue with others, through the use of qualitative data analysis software, or by creating piles of cut and pasted data. If there are multiple researchers, authors should also discuss how they worked together meaningfully in analyzing the data (this will be discussed in more detail below, in

EXERCISE 11.1



Gauging worth and rigor

- 1 In what ways is your study's topic worthy? Consider issues of:
 - a theoretical relevance
 - b practical application
 - c opportunity for social transformation
- 2 In what ways is your study interesting? How does it solve a problem or puzzle? How does it provide something new and surprising?
- 3 Is your study sufficiently rigorous? Given the topic and the contribution you hope to make, what other things might you do to ensure due diligence in terms of:
 - a collecting appropriate and sufficient data?
 - b spending enough time in the field?
 - c adopting appropriate data collection and analysis practices?

relation to inter-coder reliability). As you are finishing up data analysis and have reflected on Exercise 11.1, it makes sense to pause and ask yourself some questions about your emerging study's worth and rigor.

Sincerity

A third marker of qualitative quality is **sincerity**: this means that good qualitative research is genuine and vulnerable. Researchers share their goals, hopes, and mistakes, and they discuss how these backstage issues have implicated the fieldwork, the participants, and the data analysis. Vulnerability demonstrates openness to the life experiences of others, as well as a willingness to share aspects of your own experiences. Sincere researchers are approachable rather than self-important and friendly rather than snobbish. They consider not only their own needs but also those of their participants, readers, co-authors, and potential audiences. Sincere researchers are honest, kind, and self-deprecating. They foster sincerity through two practices discussed below: self-reflexivity and transparency.

Self-reflexivity

Self-reflexivity, as discussed throughout this book, is an honest and authentic awareness of one's own identity and research approach, and an attitude of respect for participants, audience members, and other research stakeholders. Practices of self-reflexivity include sharing one's motivations to conduct a certain study and engaging in practices that promote self-awareness and exposure. By sharing these practices, readers can feel assured that researchers have considered their role and impact in the scene.

Self-reflexivity encourages writers to be frank about their strengths and shortcomings, substantiating “their interpretations and findings with a reflexive account of themselves and the process of their research” (Altheide & Johnson, 1994, p. 489). Self-reflexive practices inform all stages of a project, beginning with heightened awareness in the early stages of the research design and progressing to later stages of fieldwork, analysis, and writing. Self-reflexive researchers consider how their bodies and intentions impact the types of data, relationships, and trust available to them (González, 2000).

Several practices are associated with the “doing” of self-reflexivity. Self-reflexive researchers make notes about others’ reactions to them. They also include themselves in the write-up of the research. Using the first person voice (e.g. “I said,” or “They reacted to me by...”) is not only allowed, but encouraged. Using the first person, “I,” reminds the reader of the researcher’s presence and influence. In contrast to mainstream journalism, self-reflexive ethnography requires explicating the *process* or *way* of knowing and *how* claims were developed (Altheide & Johnson, 1994).

How much disclosure about self-reflexivity is enough? Like a good spice, a small dose of self-reflexivity in a published research report can go a long way. Geertz (1973) made the case that he was “never [...] impressed by the argument that, as complete objectivity is impossible in these matters (as, of course, it is), one might as well let one’s sentiments run loose” (p. 29). Too much inward autobiographical detail about complaints, anxieties, wishes, or dreams can flood and overwhelm the text – unless, of course, the point of the text is to document these issues. In most research reports, it makes sense to “recognize our connections and write about them, but mainly as these connections further illuminate the reader’s understanding of the cultural event, place or practice” (Krizek, 2003, p. 149). Through a balanced approach, authors can include themselves in the scene, but not so much that it squeezes out other important objects of study (Denzin, 1997).

Transparency

Another key part of sincerity is being honest and open about the activities by which the research transpired – a feature called **transparency**. If you were only able to obtain access to a scene because your parent works there, then, to be transparent, you should share this fact with your audience. If you had to “re-create” half of the fieldnotes because a computer glitch deleted them, it is not transparent to lump the description of this fieldnote data with the rest. In short, transparency demands that research processes – which may include interactions within the context, the methodological design, analysis practices, and relationships with participants – should be self-critically and openly delineated (Altheide & Johnson, 1994; Seale, 1999).

Other issues to consider in terms of transparency are the type of field role participation, fieldnote practices, and the level of detail in transcription. Transparency suggests that authors reveal mistakes or surprises and explain how research goals and questions changed due to external constraints or unexpected challenges in the field. Readers should also know whether the research was funded, by whom, and how/if such funding shaped or determined the research design or analysis. Finally, transparency demands being up front about other people’s role in the research and acknowledging help from colleagues, participants, or student assistants. Appreciating others and acknowledging the limits of one’s own role through transparent processes such as these are certainly key practices of sincerity – something pictured in Figure 11.1 (in Researcher’s Notepad 11.1).

RESEARCHER'S NOTEPAD 11.1



Sincerity word cloud



Figure 11.1 Sincerity is a key characteristic of qualitative quality; it is made up of a number of intersecting practices, as pictured in this word word cloud that I created at www.wordle.net

Credibility

Credibility, a fourth marker for qualitative quality, is a common term that people often use without any clear definition. For example, public speakers learn to “establish credibility” by sharing their expertise or research on a topic and by persuading the audience that they are believable. For our purposes, **credibility** refers to dependability, trustworthiness (Lincoln & Guba, 1985), and expressing a reality that is plausible or *seems* true (Tracy, 2010). Good ethnography provides “a credible account of a cultural, social, individual, or communal sense of the ‘real’” (Richardson, 2000a, p. 254). If a report is credible, readers feel confident in using its data and findings to act and make decisions. Qualitative credibility is achieved through

- thick description;
- triangulation or crystallization;
- multivocality and partiality;
- engaging in member reflections with participants.

Thick description

Thick description is achieved by explicating contextual meanings specific to the cultural group at hand (Geertz, 1973), and by providing lush material details about people, processes, and activities (Bochner, 2000). Such detail gives a complex and expansionistic depiction. In qualitative research “things get bigger, not smaller and tighter, as we understand them” (González, 2000, p. 629).

Related to thick description and concrete detail is the ability of qualitative research to tap into **tacit knowledge**, which is considered to be the body of implicit meanings floating

just below the surface. In other words, tacit knowledge is the “largely unarticulated, contextual understanding that is often manifested in nods, silences, humor, and naughty nuances” (Altheide & Johnson, 1994, p. 492). If you think of an iceberg (like the one responsible for sinking the Titanic), all you may see is its top tip, sticking out of the water. This tip is akin to explicit and visible knowledge. However, the largest and most powerful part of an iceberg lies underneath, covered by water. This huge base is like tacit knowledge. Reaching tacit knowledge requires digging below the surface, to understand the importance of what is not said and how implicit core values of the group are driving action – even, or especially, when these norms cannot be easily articulated.

In order to recognize and get tacit knowledge, researchers need to spend time in the context and acquire experience of it. Just as an experienced sea captain can appreciate the difference between an iceberg and sea ice that is just skimming the surface, a researcher who has been in the field for a long time can begin to recognize the places to probe for tacit knowledge. This process includes examining the *absence* of talk and activity. Additionally, engaging in comparative research in a different but related scene is a poignant way to understand the assumed values of a certain group, system, or organization.

Crystallization/triangulation

Gathering multiple types of data seen through multiple lenses is another key way to achieve credibility. In short, findings are stronger when researchers gather their data through several sampling strategies, use more than one investigator in the field, engage multiple theoretical positions in data analysis, or use contrasting methods of data collection. This practice was originally called **triangulation**, refers to using multiple points in geographical navigation. Imagine trying to locate a museum on a map. Knowing its street name is a fine start, but additional data points – the cross-street, the postal code, or how far away it is from your hotel – can make it even easier to find it. The concept of triangulation was born in realist paradigms that aimed at ridding research of bias and at finding convergence on a single reality (or point on the map). The idea was that, when many data points converge, the findings are more credible (Denzin, 1978). One practice often associated with triangulation is that of **inter-coder reliability**, discussed in Tips and Tools 11.2.

Inter-coder reliability is only desirable when the researchers are claiming to code the data similarly. That said, there may be good reasons for two researchers to code data quite differently on purpose. Indeed multiple data points, theoretical constructs, or different researchers may appropriately come up with results that are different from one another rather than convergent. Does this mean that the research is not credible? Not necessarily. Data analyzed and gathered by two different researchers may differ because of the researchers’ age, race, gender, or past experience, and both viewpoints could shed important insight. Likewise, researchers may find one organizational value espoused in some data sources, such as interviews and training documents (e.g. that good employees tell the truth). Meanwhile they may discover in participant observation that “good” employees regularly fudge the numbers or tell white lies. Findings from both data texts may be equally “true” and show the complexities of the scene.

Indeed making use of multiple data points and researcher points of view, even when they do not converge, is still a practice toward qualitative credibility. The notion of **crystallization** refers to such a practice while avoiding the realism associated with the term “triangulation.” The multiple facets of crystals “reflect externalities *and* refract within themselves, creating different colors, patterns, and arrays, casting off in different directions. What we see

TIPS AND TOOLS 11.2



Inter-coder reliability

Although the overall concept of reliability is not usually applicable to qualitative research (as discussed in the opening of this chapter), inter-coder reliability is important when a team of qualitative researchers want to ensure they are coding and classifying data from a single study in a similar way.

In such cases, they engage in practices to ensure inter-coder reliability. Inter-coder reliability can be calculated in a variety of manners, the most common of which are percent agreement, Scott's pi (p), Cohen's kappa (k), and Krippendorff's alpha (α) (see Lombard, Snyder-Duch, & Bracken, 2002, for a review). Different disciplines and journals have varying expectations for computing inter-coder reliability. In what follows I describe how to engage in inter-coder reliability via percent agreement:

- 1 Through dialogue and consultation, collaborators create a common coding scheme.
- 2 Collaborators work together coding data to try to become consistent in their understandings.
- 3 Collaborators separate and, working independently, analyze the same subset (usually at least 10%).
- 4 Collaborators come back together to compare their coding and to compute inter-coder reliability.
 - a This is calculated by taking the number of codes that the researchers agreed upon (e.g. 9) and dividing it by the total number of pieces of data coded (e.g. 10).
 - b The higher the agreement rate (9/10), the more reliable (or consistent), the analysis. An agreement rate of 90 percent or higher is considered appropriate by post-positivist qualitative researchers (Miles & Huberman, 1994).
- 5 When collaborators reach an appropriate agreement rate, they may then assume that they are coding the data similarly, and therefore they can break up the rest of the data and analyze them independently.

depends upon our angle of repose" (Richardson, 2000b, p. 934). Through crystallization, as this notion was developed by Ellingson (2008), researchers are encouraged to engage in multiple types of data collection, at multiple points in time, with multiple co-researchers, in order to construct a multi-faceted, more complicated, and therefore more credible picture of the context.

Multivocality

One way of practicing crystallization is through **multivocality** – the inclusion of multiple voices. This means analyzing social action from a variety of participants' points of view and highlighting divergent or disagreeable standpoints. Multivocality also requires that authors be self-aware of how their own and their participants' subjectivities vary in the field – in terms of race, gender, age, education, class, or sexuality. Credibility is enhanced by considering how these differences play a role in conflicting intentions or in narratives of contextual practices and performances. For instance, different groups may have very different explanations and assessments of humor. Some may find a joke harmless and fun, while others may find it mean-spirited and divisive.

Writing a multivocal analysis can be facilitated through collaboration with research participants. As described in Chapter 3, participatory, action, autoethnographic, and feminist approaches seek out and include participant input along the way. By taking all these factors in, researchers can better ensure that their voices are represented in multifaceted ways, nuanced, and ultimately more credible in the final report (Ellis, 2007).

Member reflections

In relation to multivocality, researchers can also include participants in the analysis of data and findings. I use the phrase **member reflections** to refer to occasions that “allow for sharing and dialoguing with participants about the study’s findings, providing opportunities for questions, critique, feedback, affirmation and even collaboration” (Tracy, 2010, p. 844). Such a practice includes sitting down with your participants and sharing in-process analyses or conclusions, making note of their reactions, and including these reactions in further cycles of data analysis.

Member reflections are different from *member checks*, *member validation*, and *host verification* (Lindlof & Taylor, 2011) – all practices that emphasize the need for *correspondence* between the researcher’s findings and the participants’ viewpoints. Rather, member reflections suggest that participant feedback is valuable not as a measure of validity, but as a space for additional insight and credibility. Through the collaboration and elaboration that occur as a result of member reflections, new data are produced that “throw fresh light on the investigation and [...] provide a spur for deeper and richer analyses” (Bloor, 2001, p. 395).

Through member reflections, researchers may also appreciate the extent to which their findings are understandable and meaningful to the participants, themselves. In the reflection process participants can react, agree, or point out problems with the analysis. Providing opportunities for member reflections is not only ethical – especially when participants have dedicated significant patience, time, resources, and energy to the project – but also speaks volumes about the study’s credibility.

What if members disagree with the findings, or dislike their own portrayal in them? Does this mean the research is less credible or valid? Perhaps, but not necessarily. Authors must be comfortable with critique and disagreement from a range of audiences, especially if they subscribe to paradigms that view the world as contested and constructed. Good researchers pay attention to the variety of reactions received, incorporating them as they continue to gather, analyze, and write. At the same time, they should not automatically change the direction of their analysis because participants disagree; “members’ responses to researchers’ accounts are provisional and subject to change” (Bloor, 2001, p. 391). Depending on the season, time of day, or context, they may protest at one point but certify findings at another. Researchers should create options for input and consider it as they analyze the data and write up the final report.

Resonance

A fifth key marker of qualitative quality is **resonance**, considered to be the feature of the text that meaningfully reverberates and impacts an audience. Many people erroneously assume that formal generalizability is the only way to achieve resonance. Most qualitative researchers, rather than trying to prove the generalizability of research

through statistics, ensure that their research resonates by choosing specifically revealing cases or contexts of study. As Flyvbjerg notes:

One can often generalize on the basis of a single case, and the case study may be central to scientific development via generalization as supplement or alternative to other methods. But formal generalization is overvalued as a source of scientific development, whereas ‘the force of example’ and transferability are underestimated. (Flyvbjerg, 2011, p. 305)

Indeed, as I discuss here, resonance can be achieved through several different practices.

Transferability and naturalistic generalization

When readers intuitively believe that research findings correspond to something significant in their own world, then resonance has been accomplished through **transferability**. For instance, if readers of my cruise ship research relate its emotional labor and burnout findings to their own situation (e.g. to their work in a restaurant or in a theme park), they are *transferring* the findings. Transferability is different from formal generalizations, in which the researcher engages in randomized sampling and “objective” scientific practices to generate context-free and formally generalizable knowledge.

Qualitative researchers also achieve resonance by helping readers feel as if they have been there. The concept of **naturalistic generalization** (Stake & Trumbull, 1982) refers to this process – in which readers appreciate a study’s findings and then intuitively apply them to their own situations. For example, Geertz (1973) claims that Balinese cockfighting practices are symbolic of the culture’s overall concerns with violence and status. These claims rest upon data that Geertz describes so thickly and richly that readers can viscerally feel and understand how the implications of the data might transfer – or naturalistically generalize – to their own culture (e.g. how sports such as football symbolize violence and status in US culture).

Naturalistic generalizations can actually have more impact than statistical generalizations, because participants feel as though they have made their own useful applications rather than been told by the author what meanings to accept. Indeed, Stake and Trumbull argue that, compared to formal knowledge, the readers’ “practice is guided far more by personal knowings, based on and gleaned from personal experience” (p. 5). So how might researchers write their report so as to make it transferable and naturalistically generalized? As I explain below, a qualitative researcher can communicate the impact of her findings by writing so aesthetically that the reader can imagine and personally transfer these findings to a range of familiar contexts.

Aesthetic merit

Another means toward resonance is that of writing a text with **aesthetic merit** – in other words making it imaginative, artistic, beautifully written, and capable of emotionally affecting the reader (Goodall, 2008). Have you ever read something so moving that you laughed out loud, cried, felt sick to your stomach, or felt inspired to change the world? If so, then the text has aesthetic merit. It engaged your feeling and interpretive response. It was not boring.

For qualitative researchers, writing an excellent research report requires not just the basic qualities of clarity and organization. Aesthetic writers use a variety of literary and

evocative styles, including personal narrative, storytelling, and emotional approaches that help readers tap into their own bodies (Holman Jones, 2005). Aesthetic texts are interactive, descriptive, and evocative (Scarduzio, Giannini, & Geist-Martin, 2011) – they move the “heart and the belly” as well as the “head” (Bochner, 2000, p. 271).

One such text is Ronai’s (1992) layered account of exotic dancing. Carol Rambo-Ronai returned to her former job as an erotic dancer, conducting an autoethnography. She made use of Ellis’s (1991) emotional sociology and called upon her own emotional experience as a method to describe, examine, and theorize. Ronai’s raw emotions in the field helped her to understand and explain how stripping led simultaneously to feelings of power and powerlessness, repulsion and superiority. Readers feel what Ronai feels as one client paws at her on stage, another attempts unsuccessfully to stick his tongue down her throat, and another cat-calls her in the parking lot. Her writing invites readers to see, taste, touch, and smell that men’s club, complete with a nicotine sheen on its walls. Through a conscious self-examination of her felt emotions, Ronai vividly shows how erotic dancing impacts private and social experiences. Certainly she succeeds in terms of Richardson’s (2000a) evaluative question about good qualitative research: Does this affect me emotionally or intellectually?

A focus on aesthetics does not preclude rigorous scientific practice. Rather, it means that researchers understand that the way they creatively construct their literary tale impacts the resonance of their research report. Richardson (2000a) explains that “creative arts is one lens through which to view the world; analytical/science is another.” She says that “we see best with both lenses focused and magnified” (p. 254). Aesthetic merit opens a beautiful path toward achieving resonance by provoking vicarious emotional experience in the reader (Ellis, 1995).

Significant contribution

We have now covered five key characteristics of qualitative quality. A sixth – which can make or break publication in academic journals – is the significance of the study’s contribution. Significance is largely judged by whether the findings extend, transform, or complicate a body of knowledge, theory, or practice in new and important ways. In short, significant research serves to “bring clarity to confusion, make visible what is hidden or inappropriately ignored, and generate a sense of insight and deepened understanding” (Tracy, 1995, p. 209).

This insight or deepened understanding need not be huge, but it must impact the current knowledge landscape. Making a significant contribution requires familiarity with the broad literature, delving deeper into a particular issue, reading and learning as much as possible about that issue, finding its boundaries, and then pushing those limits to see how the area might benefit from more research. (Matt Might, <http://matt.might.net/> provides a captivating visual metaphor in his illustrated guide to getting a PhD <http://matt.might.net/articles/phd-school-in-pictures/>.) After all this reading and boundary pressing, researchers engage in their own original research, with the hope that they may press the knowledge boundaries a little bit further. That little incremental addition, or dent in the knowledge boundary, is a significant contribution. It’s a gift of your research to the body of knowledge – one that readers can learn from and use in their own future research.

Theoretically significant research extends, builds, or critiques disciplinary knowledge, helping to explain social life in unique ways. At its most basic level, theoretical significance

may come in the form of applying an established theory in a new context. For example, Baxter (1990) identified three key dialectical tensions to help explain the development of romantic relationships. Since that time, dialectical theory has been applied hundreds of times to other settings, to explain relational contradictions amongst friends, family members, and co-workers. Applying existing theory to a new context is usually an adequate contribution in undergraduate student papers or theses. **Conceptual development**, however, goes a step further: research *builds* theory beyond the existing literature and offers new and unique understandings. Conceptual development is more difficult than simply applying existing theory to a new setting. Yet achieving the former is required in most graduate-level theses and dissertations, as well as in scholarly publications.

For instance, let us consider a study of burnout in a group of professionals (say, among attorneys). The researcher might begin with the original conceptualization of burnout that emanated from Maslach's studies with social service professionals (Maslach, 1976). Using the data gathered from attorneys, the researcher could discuss that burnout among professionals has some similarities with Maslach's notions of burnout, but is also somewhat different. Through the research, the authors may suggest that another concept – say “tedium” – better captures the unique stress and burnout that occurs among professionals (Pines, Aronson, & Kafry, 1981). This study builds upon past research burnout, but it offers a new concept of tedium, which is potentially useful to future researchers.

Speaking of future researchers, **heuristic significance** is the quality of research that prompts curiosity in others, moving them to act, perform additional investigations, or examine how the concept might play out in a different context or group. As an example of heuristically significant research, we might examine Lutgen-Sandvik's (2003) cycle of workplace abuse. Once this model existed, other communication scholars became interested in issues of workplace bullying, civility, and dignity – and this later research investigated abuse in a variety of workplace settings, looping back and extending Lutgen-Sandvik's model. Likewise, one can find out that the concept of compassion has great heuristic significance; and one reaches this insight by tracing its history, from compassion fatigue (Figley, 1995) to a process of noticing, feeling, and responding (Kanov et al., 2004), to a process of noticing, connecting, and responding (Miller, 2007), and then to a process of recognizing, relating, and (re)acting (Way & Tracy, 2012). Across these various studies the concept has changed and progressed, hopefully in a way that more precisely and richly explains the practice of compassion at work.

One way authors may bolster heuristic significance is by specifically discussing new directions or questions for research, suggesting what we still do not know and how researchers might attend to such issues in subsequent studies. Furthermore, writing about the research in an engaging and accessible manner may prompt an entire range of potential audiences, including lay people and policy makers, to act upon the findings.

Speaking of affecting a range of audiences, research may also offer **practically significant research** contributions through helpful and useful insight in the day-to-day life of key stakeholders. The phronetic, problem-based contextual approach described in this book is specifically designed to result in findings of practical significance. For example, on the basis of her ethnographic study of the ways employees manufactured a mood of innovation, Elizabeth Eger (formerly Rush) (Rush, 2010) offers a number of recommendations for mood interventions that are more fulfilling and energizing than the exhausting and depleting practices she witnessed at her research site. Such research certainly succeeds in enabling “the training or calibration of human judgment” (Schwandt, 1996, p. 69) and “the capacity for practical wisdom” (p. 70).

EXERCISE 11.2



Gauging significance

- 1 In what ways will your study extend, complicate, or build theoretical knowledge? How about methodological practice?
- 2 In what ways might your study lead to heuristic significance, encouraging future researchers to take up this area of research and study?
- 3 In what ways does your research implicate everyday practice? What practical tips could lay people derive from your study?

Phronesis and practical wisdom are also related to the ways the research may help transform an injustice or help others learn how to replicate a liberating environment. Lather (1986) offers the notion of **catalytic validity** to specifically refer to research that provides a political consciousness that catalyzes/moves cultural members to act. Action researchers, as discussed in Chapter 3, team up with research participants to examine issues that are contextually important and to provide findings that are helpful to cultural members. Action research allows such members to work with experienced researchers in order to critique status quo problems and to cultivate notions for transformation.

Finally, **methodological significance** is achieved when methodology is approached in a new, creative, or insightful way. For instance, one might take a theory or a concept that has mostly been studied quantitatively, through experiments or self-report surveys, and study it instead through participant observation, interviews, or focus groups. New methods may not only offer fresh theoretical insight, they may sharpen and strengthen that method. One might engage in what Richardson (2000b) refers to as “creative analytic practices” – approaching insight through dance, performance, and art. Or researchers might practice alternative analysis activities – such as asking a participant to take pictures on a disposable camera, write captions for these pictures, and then talk through their meaning with the researcher. Methodological significance provides insight in terms of our craft skills associated with collecting, managing, and analyzing data, and, given the rich texture of the qualitative landscape, this is an area ripe for expansion.

Ethical research practice

The seventh characteristic I will discuss in terms of qualitative quality is that of ethics. Ethical research practice is a thread throughout this entire book, in the form of self-reflexivity, access, participation, interviewing, fieldwork, transcription, and writing. That said, ethics is so important for quality that there is good reason to highlight it separately here. Practicing ethics in qualitative research requires consideration of (a) procedural rules and procedures; (b) the specific ethics of the context we are studying; and (c) the ethics of working – sometimes quite closely and intimately – with research participants.

Procedural ethics

Procedural ethics refer to ethical actions that are prescribed by certain organizational or institutional review boards (IRB) as being universal or necessary. IRB requirements, as discussed in Chapter 5, are:

- do no harm;
- avoid deception;
- get informed consent;
- ensure privacy and confidentiality.

We have already discussed the importance of transparency and honesty. Procedural ethics likewise encompasses the importance of accuracy and of not misleading the reader through omission, exaggeration, or inappropriate attribution. Procedural ethics also refers to consent: “Weak consent usually leads to poorer data: Respondents will try to protect themselves in a mistrusted relationship, or one formed with the researcher by superiors only” (Miles & Huberman, 1994, p. 291). Creating and cultivating trust with participants is imperative. Some believe that small measures of deception are acceptable when their potential social benefits are clear (Sales & Folkman, 2000). Nonetheless, members must know that their participation is voluntary and understand how to opt out.

Participants also have a right to confidentiality. In order to protect participant identity and privacy, researchers should secure research data (for example by storing them in locked offices or on password-protected websites) and strip them of identifiers before sharing them with co-researchers, assistants, readers, or audience members. Stripping qualitative data of identifiable material can be tricky. Researchers should carefully consider how certain actors might be identified even if their name is a pseudonym: if you share a story about an “elderly boss who stole from the company,” participants may be able to quickly deduce identity if there is only one “elderly” boss. Procedural ethics about confidentiality and anonymity encourages researchers to carefully consider how they portray (or strategically conflate) sensitive data.

Situational ethics

Whereas procedural ethics provides universal edicts for all research, **situational ethics** refers to ethical issues that arise in specific contexts or sample populations. Everyone can think of acts that may be ethical in some situations or with some people, but not in other situations with other people. For instance, secretly videotaping a famous American preacher whose sermons are regularly televised has different ethical implications from secretly videotaping a medicine man celebrating an intimate ceremony in a developing world.

Situational ethics focuses on reasoned considerations about the specific situation (Fletcher, 1966); therefore it treats predetermined moral principles – such as those upheld by institutional review boards – as flexible guidelines rather than unassailable edicts. A situational ethics like utilitarianism, with its concern with “the greater good,” motivates researchers to ask whether the potential benefits of the research outweigh its costs. Likewise, researchers may consider the extent to which the study’s potential findings justify ethically questionable practices. Consider This 11.1 raises questions that urge reflection upon such issues. Certainly, there are no quick fixes here; but, as Ellis (2007) notes, a situational ethics asks that we “constantly have to consider which questions to ask, which secrets to keep, and which truths are worth telling” (p. 26).

Finally, qualitative researchers can also usefully reflect on the notion of **relational ethics**, an ethics of care that “recognizes and values mutual respect, dignity, and connectedness between researcher and researched, and between researchers and the communities in which they live and work” (Ellis, 2007, p. 4). A relational ethic means being aware of one’s own role and impact on relationships and treating participants as whole people rather than as just subjects from which to wrench a good story. Related to this, the notion of **feminist communitarianism** (Christians, 2005) suggests that researchers should collaborate with their participants, keep their promises, and put relationships and communal well-being at the top of their priorities. Consider This 11.2 provides several questions that may help prompt such an ethical moral compass.

To summarize, ethical research includes the consideration of procedural rules and regulations, as well as of situational preferences and participants’ needs. Ethical researchers vigilantly consider the impact of their practices *throughout* the inquiry. Ethical obligations are complex, and sometimes larger structural research goals and everyday micro-practices may conflict. Further, even if an action is permissible from the point of view of formal standards, if something *feels* inappropriate, then it probably is inappropriate.

Meaningful coherence

The final, and anchoring, characteristic of qualitative quality is **meaningful coherence**. In using the concept of coherence, I do *not* mean that a text cannot or should not be written in a way that is layered or intentionally jarring (this will be discussed in Chapter 12). Furthermore, meaningful coherence does not suggest that researchers cannot borrow and combine concepts from different theories. In fact, a hallmark of grounded qualitative research is the notion of novel theoretical juxtapositions and of borrowing from other fields, models, and assumptions (Tracy, 2012). Rather, by “meaningfully coherent” I mean that qualitative studies should: “(a) achieve their stated purpose; (b) accomplish what they espouse to be about; (c) use methods and representation practices that partner well with espoused theories and paradigms; and (d) attentively interconnect literature reviewed with research foci, methods, and findings” (Tracy, 2010, p. 848).

For example, if a researcher is interested in better understanding “social support,” a meaningful coherent study actually examines issues of “social support,” and not other concepts such as “venting” or “bitching.” In this way the concept is similar to **discriminant validity** – a phrase used in quantitative research to refer to the quality of a measurement device (a survey, or an experiment) to examine the specific issue intended to be studied – and not something else. Of course, in most qualitative studies the researcher, rather than a certain measuring device, is the instrument.

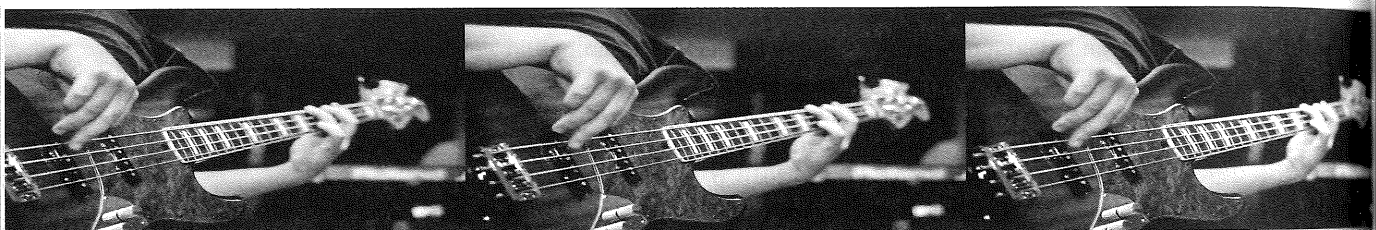
Meaningful coherence is also about the logical and intuitive connection of various arguments or concepts in a single paper. For instance, an *interpretive* theoretical approach, which suggests that meaning emerges from the voices of participants, should only invoke issues of power when or if participants bring up these issues themselves. If the researchers bring in structural issues of power themselves, without participant contribution, it would be more coherent to use a *critical* theoretical approach. Researchers must mindfully synch their stated research goals with theories, their research design, and their methodology.

Here is another example. Grounded theory as originally conceived by Glaser and Strauss (1967) is quite realist in nature: it develops an overriding story or set of themes

as grounded and “real” in any group of data. Hence, if a researcher espoused a paradigmatic framework of postmodernism – which views reality as fragmented and largely unknowable – it would be incoherent to reference initial conceptualizations of grounded methodology (Tracy, 2010). Certainly, distinct concepts emerging from grounded theory, like the constant comparative method or analytic memos, may be compatible with a postmodern analysis. However, the researcher should know enough about the assumptions of postmodern and grounded theory to realize that the adoption of each, whole-cloth, is not coherent.

Likewise, meaningful coherence requires that researchers demonstrate their understanding that certain ways of writing – such as Richardson’s (2000a) creative analytic practices – emanate from certain paradigmatic assumptions (e.g. postmodernism). Does this mean that creative analytic practices cannot be used with a realist interpretive approach? Maybe; but maybe not. At the very least, if these two approaches were to be paired together, in order to be meaningfully coherent the author would have the responsibility of explaining how and why they played well together.

Finally, meaningfully coherent studies hang together well (Fisher, 1987). The literature reviewed establishes the context for interpreting the findings. Research questions or purposes arise logically from the literature. The goals are achieved in the analysis, and the conclusions and implications speak to issues, questions, or controversies in the literature. Incoherent studies, in contrast, may open with one literature, but have findings and implications that relate to another. In sum, after reading a meaningfully coherent study, readers should clearly understand the purpose of the piece and feel as though its findings were delivered in relation to its stated goals.



FOLLOWING, FORGETTING, AND IMPROVISING

This chapter has presented a framework for qualitative quality. However, as researchers, we sometimes may fall short, deviate, forget, or purposefully improvise. In some cases our human instrument shows its innate humanness by not being able to achieve everything all of the time.

You are not alone if you feel that trying to meet one of these quality criteria makes it difficult to reach another. Researchers are often faced with a choice between two goals, such as validity versus avoiding harm, scientific understanding versus individual rights, detached inquiry versus help, help-giving versus confidentiality, and freedom of inquiry versus political advantage (Miles & Huberman, 1994). Qualitative researchers must consistently