Copyright 2016. Jossey-Bass.
All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

Chapter Nine

Dealing with Validity, Reliability, and Ethics

All research is concerned with producing valid and reliable knowledge in an ethical manner. Being able to trust research results is especially important to professionals in applied fields because practitioners intervene in people's lives. No classroom teacher, for example, will want to experiment with a new way of teaching reading, nor will a counselor want to implement a new technique to engage with a bereaved family without some confidence in its probable success. But how can you know when research results are trustworthy? They are trustworthy to the extent that there has been some rigor in carrying out the study. Because qualitative research is based on assumptions about reality different from those of quantitative research (see Chapter One), the standards for rigor in qualitative research necessarily differ from those of quantitative research. However, since both the criteria and the terminology for discussing and assessing rigor in qualitative research are in flux (Denzin & Lincoln, 2011; Lichtman, 2013), we have chosen to discuss trustworthiness and rigor in interpretive qualitative research with reference to the traditional terminology of validity and reliability, though we recognize these are contested terms.

Ensuring validity and reliability in qualitative research involves conducting the investigation in an ethical manner. Although well-established guidelines for the ethical conduct of research date back to the late 1940s, only within the last few decades has attention been given to the ethical concerns unique to qualitative research. We conclude the chapter by considering how ethical practices are also important in establishing the trustworthiness of your study.

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE
AN: 1022562 ; Sharan B. Merriam, Elizabeth J. Tisdell.; Qualitative Research : A Guide to Design and Implementation
Account: s6222068.main.eds

Validity and Reliability

To have any effect on either the practice or the theory of a field, research studies must be rigorously conducted; they need to present insights and conclusions that ring true to readers, practitioners, and other researchers. The applied nature of most social science inquiry thus makes it imperative that researchers and others have confidence in the conduct of the investigation and in the results of any particular study. Lincoln, Lynham, and Guba (2011, p. 120) underscore this point by asking whether a study's findings are "sufficiently authentic . . . that I may trust myself in acting on their implications? More to the point, would I feel sufficiently secure about these findings to construct social policy or legislation based on them?"

Regardless of the type of research, validity and reliability are concerns that can be approached through careful attention to a study's conceptualization and the way in which the data are collected, analyzed, and interpreted, and the way in which the findings are presented. Firestone (1987) explores how the quantitative and qualitative paradigms employ different rhetoric to persuade consumers of their trustworthiness. "The quantitative study must convince the reader that procedures have been followed faithfully because very little concrete description of what anyone does is provided. The qualitative study provides the reader with a depiction in enough detail to show that the author's conclusion 'makes sense'" (p. 19). Further, "the quantitative study portrays a world of variables and static states. By contrast the qualitative study describes people acting in events" (p. 19). In the more recent mixed methods designs, both qualitative and quantitative criteria are applied to assess the trustworthiness of the study (Creswell, 2015).

Research designs are based on different assumptions about what is being investigated, and they seek to answer different questions. If, as in the case of qualitative research, *understanding* is the primary rationale for the investigation, the criteria for trusting the study are going to be different than if discovery of a law or testing a hypothesis is the study's objective. What makes experimental studies scientific or rigorous or trustworthy is the researcher's careful design of the study, applying standards well developed and accepted by the scientific community. Qualitative research also has strategies for establishing the authenticity and trustworthiness of a study—strategies based on worldviews and

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

questions congruent with the philosophical assumptions underlying this perspective (see Chapter One).

Many writers on the topic argue that qualitative research, which is based on different assumptions about reality and different worldviews, should consider validity and reliability from a perspective congruent with the philosophical assumptions underlying the paradigm. This may even result in naming the concepts themselves differently, as Lincoln and Guba (1985) did. *Credibility, transferability, dependability,* and *confirmability*—as substitutes for *internal validity, external validity, reliability,* and *objectivity*—were for a while widely adopted in qualitative research. More recent writing from postmodern, poststructural, constructivist, critical, and action research perspectives (Cho & Trent, 2006; Denzin & Lincoln, 2011; Herr & Anderson, 2015; Patton, 2015; Richardson & St. Pierre, 2005) calls for the careful thinking through of totally different conceptualizations of validity and reliability. Denzin and Lincoln (2000), for example, consider the postmodern turn in qualitative research as problematic for evaluating qualitative research. "This is the legitimation crisis. It involves a serious rethinking of such terms as *validity, generalizability,* and *reliability,* terms already retheorized" in other types of qualitative research (p. 17, emphasis in original). More recently, Lincoln, Lynham, and Guba (2011) proposed two forms of rigor—methodological, related to the application of methods, and interpretive, related to judging outcomes, that is, "Can our co-created constructions be trusted to provide some purchase on some important human interpretation?" (p. 121).

Lichtman (2013) uses a continuum to capture this fluidity in changing notions of defining and assessing trustworthiness in qualitative research. Prior to 1990 the concepts of objectivity, reliability, and internal validity were used to assess qualitative research. In the next decade, 1990–2000, the concepts of *credibility, transferability, dependability,* and *confirmability* (Guba & Lincoln, 1981; Lincoln & Guba, 1985) were thought to be more suitable criteria. Beginning in 2000 she identifies both "a resurgence of interest" in traditional criteria and criteria that represent "differing points of view. These criteria tend to emphasize the role of the researcher, for example," and they are "very much influenced by some of the newer ideas of post-structuralism, feminism, and postmodernism. Politics and power also play a critical role here" (p. 292).

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

Furthermore, with the wide variety of types of qualitative research (see Chapters Two and Three), there are bound to be differences in criteria for validity and reliability. Creswell (2013), for example, applies somewhat different criteria for evaluating how "good" a narrative study is compared to phenomenological research, grounded theory research, ethnographic research, or case study research. In a narrative study he suggests that good narrative tells an engaging story versus one criterion of a good ethnography being "a detailed description of the cultural group" (p. 263). Lichtman (2013) offers her own "personal criteria" for "a good piece of qualitative research" (p. 294). These include being explicit about the researcher's role and his or her relationship to those studied, making a case that the topic of the study is important, being clear about how the study was done, and making a convincing presentation of the findings of the study.

Similar to Lichtman's "personal criteria" is Tracy's (2013) "big-tent" criteria for conducting "excellent" qualitative research. Her eight criteria are that the research (1) be on a worthy topic; that it be conducted with (2) rich rigor and (3) sincerity—that is, transparency of methods—and (4) credibility; that the research (5) resonates with a variety of audiences and (6) makes a significant contribution; (7) that it attends to ethical considerations; and finally, (8) that the study have meaningful coherence; that is, "meaningfully interconnects literature, research, questions/foci, findings, and interpretations with each other" (p. 230). Wolcott (1994) takes yet another direction, arguing "the absurdity of validity" (p. 364). Instead of validity, what he seeks "is something else, a quality that points more to identifying critical elements and wringing plausible interpretations from them, something one can pursue without becoming obsessed with finding the right or ultimate answer, the correct version, the Truth" (pp. 366–367). For Wolcott that "something else" is understanding.

To further underscore the complexity of addressing the issue of validity and reliability in a world of burgeoning qualitative research designs, Patton (2015) offers *seven* "alternative sets of criteria for judging the quality and credibility of qualitative inquiry" (p. 680). Depending upon the type of research, he suggests criteria for (1) traditional scientific, (2) constructivist, (3) artistic, (4) systems/complexity, (5) participatory, (6) critical, and (7) pragmatic/utilization focused research.

Those conducting qualitative investigations do not want to wait for the research community to develop a consensus as to the appropriate criteria for assessing validity and reliability, if indeed that is even possible. While the theoretical debate goes on, there are immediate needs to be met in the field. As Stake (2005) notes, knowledge gained in an investigation "faces hazardous passage from writing to reading. The writer seeks ways of safeguarding the trip" (p. 455). Further, qualitative researchers need to respond to the concerns of outsiders, many of whom may be unfamiliar with or blatantly challenging of the credibility of qualitative research. Exhibit 9.1, for example, is a list of sample questions often asked of qualitative researchers. Each question asks something about the validity or reliability of qualitative research.

Exhibit 9.1. Challenging the Trustworthiness of Qualitative Research.

1. What can you possibly tell from an *n* of 1 (3, 15, 29, and so on)?
2. What is it worth just to get the researcher's interpretation of the participant's interpretation of what is going on?
3. How can you generalize from a small, nonrandom sample?
4. If the researcher is the primary instrument for data collection and analysis, how can we be sure the researcher is a valid and reliable instrument?
5. How will you know when to stop collecting data?
6. Isn't the researcher biased and just finding out what he or she expects to find?
7. Without hypotheses, how will you know what you're looking for?
8. Doesn't the researcher's presence result in a change in participants' normal behavior, thus contaminating the data?
9. Don't people often lie to field researchers?
10. If somebody else did this study, would they get the same results?

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

Fortunately, several strategies can be used to enhance the validity and reliability of qualitative studies. In keeping with our goal of introducing qualitative research to our readers based upon a constructivist worldview, we have chosen to focus on methodological rigor; that is, what you, as a researcher can do to ensure trustworthiness in your study. The following sections address the specific concerns in constructivist qualitative research with respect to internal validity, reliability, and external validity—or what Lincoln and Guba (1985) call credibility, consistency/dependability, and transferability—and suggest appropriate strategies for dealing with each of these issues.

Internal Validity or Credibility

Internal validity deals with the question of how research findings match reality. How congruent are the findings with reality? Do the findings capture what is really there? Are investigators observing or measuring what they think they are measuring? Internal validity in all research thus hinges on the meaning of reality. Becker (1993) humorously points out that "reality is what we choose not to question at the moment," and "the leading cause of stress amongst those in touch with it" (p. 220). On a more serious note, Ratcliffe (1983) offers an interesting perspective on assessing validity in every kind of research. It should be remembered, he suggests, that (1) "data do not speak for themselves; there is always an interpreter, or a translator" (p. 149); (2) that "one cannot observe or measure a phenomenon/event without changing it, even in physics where reality is no longer considered to be single-faceted"; and (3) that numbers, equations, and words "are all abstract, symbolic representations of reality, but not reality itself" (p. 150). Validity, then, must be assessed in terms of something other than reality itself (which can never be grasped). That "something other than reality itself" is Lincoln and Guba's (1985) notion of credibility; that is, are the findings *credible*, given the data presented?

One of the assumptions underlying qualitative research is that reality is holistic, multidimensional, and ever-changing; it is not a single, fixed, objective phenomenon waiting to be discovered, observed, and measured as in quantitative research. Assessing the isomorphism between data collected and the "reality" from

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

which they were derived is thus an inappropriate determinant of validity. In writing about his scientific journey to the Sea of Cortez more than seventy years ago, Steinbeck (1941) eloquently contrasted the two views of reality:

> The Mexican sierra has "XVII–15–1X" spines in the dorsal fin. These can easily be counted. But if the sierra strikes hard on the line so that our hands are burned, if the fish sounds and nearly escapes and finally comes in over the rail, his colors pulsing and his tail beating the air, a whole new relational externality has come into being—an entity which is more than the sum of the fish plus the fisherman. The only way to count the spines of the sierra unaffected by this second relational reality is to sit in a laboratory, open an evil smelling jar, remove a stiff colorless fish from formalin solution, count the spines, and write the truth "D. XVII–15–1X." There you have recorded a reality which cannot be assailed—probably the least important reality concerning either the fish or yourself. The man with his pickled fish has set down one truth and has recorded in his experience many lies. The fish is not that color, that texture, that dead, nor does he smell that way. (p. 2)

Maxwell (2013) concurs that one can never really capture reality. "Validity is never something that can be proved or taken for granted. Validity is also relative: It has to be assessed in relationship to the purposes and circumstances of the research, rather than being a context-independent property of methods or conclusions" (p. 121).

Then what *is* being studied in qualitative research, and how does a researcher assess the validity of those observations? What is being investigated are people's constructions of reality—how they understand the world. And just as there will be multiple accounts of eyewitnesses to a crime, so too there will be multiple constructions of how people have experienced a particular phenomenon, how they have made meaning of their lives, or how they have come to understand certain processes.

Because human beings are the primary instrument of data collection and analysis in qualitative research, interpretations of reality are accessed directly through their observations and interviews. We are thus "closer" to reality than if a data collection

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

instrument had been interjected between us and the participants. Most agree that when rigor is viewed in this manner, internal validity is a definite strength of qualitative research. In this type of research it is important to understand the perspectives of those involved in the phenomenon of interest, to uncover the complexity of human behavior in a contextual framework, and to present a holistic interpretation of what is happening.

LeCompte and Preissle (1993) list four factors that lend support to the claim of high internal validity of ethnographic research:

> First, the ethnographer's common practice of living among participants and collecting data for long periods provides opportunities for continual data analysis and comparison to refine constructs; it ensures a match between researcher categories and participant realities. Second, informant interviews, a major ethnographic data source, are phrased in the empirical categories of participants; they are less abstract than many instruments used in other research designs. Third, participant observation—the ethnographer's second key source of data—is conducted in natural settings reflecting the life experiences of participants more accurately than do more contrived or laboratory settings. Finally, ethnographic analysis incorporates researcher reflection, introspection, and self-monitoring that Erickson (1973) calls disciplined subjectivity, and these expose all phases of the research to continual questioning and reevaluation. (p. 342)

Though qualitative researchers can never capture an objective "truth" or "reality," there are a number of strategies that you as a qualitative researcher can use to increase the "credibility" of your findings, or as Wolcott (2005, p. 160) writes, increase "the correspondence between research and the real world." Probably the best-known strategy to shore up the internal validity of a study is what is known as *triangulation.* Usually associated with navigation or land surveying, wherein two or three measurement points enable convergence on a site, the best-known discussion of triangulation is Denzin's (1978), in which he proposes four types: the use of multiple methods, multiple sources of data, multiple investigators, or multiple theories to confirm emerging findings. The use of multiple theories such as approaching "data with several

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

hypotheses in mind, to see how each fares in relation to the data" (Seale, 1999, p. 54) is less common in qualitative research than are the other three forms.

With regard to the use of multiple *methods* of data collection, for example, what someone tells you in an interview can be checked against what you observe on site or what you read about in documents relevant to the phenomenon of interest. You have thus employed triangulation by using three methods of data collection—interviews, observations, and documents.

Triangulation using multiple sources of *data* means comparing and cross-checking data collected through observations at different times or in different places, or interview data collected from people with different perspectives or from follow-up interviews with the same people. *Investigator* triangulation occurs when there are multiple investigators collecting and analyzing data. Patton (2015, p. 665) suggests a related strategy, that of "*triangulating analysts*—that is, having two or more persons independently analyze the same qualitative data and compare their findings" (emphasis in original). This notion of multiple researchers has also been discussed in other contexts as collaborative or team research. In participatory research, where the goal of the research is political empowerment, the participants along with the researcher collectively define the problem to be addressed, conduct the study, and engage in collective action to bring about change.

Thus, *triangulation*—whether you make use of more than one data collection method, multiple sources of data, multiple investigators, or multiple theories—is a powerful strategy for increasing the credibility or internal validity of your research. As Patton (2015) explains, "triangulation, in whatever form, increases credibility and quality by countering the concern (or accusation) that a study's findings are simply an artifact of a single method, a single source, or a single investigator's blinders" (p. 674).

It might be noted that as with other strategies for ensuring trustworthiness in qualitative research, triangulation is being revisited in the literature from a postmodern perspective. Richardson (2000; see also Richardson & St. Pierre, 2005) points out that triangulation assumes a "'fixed point' or 'object' that can be triangulated." But in postmodern research, "we do not triangulate; we *crystallize.* We recognize that there are far more than three sides

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

from which to approach the world" (Richardson, 2000, p. 934). Crystals exhibit "an infinite variety of shapes, substances, transmutations, multidimensionalities, and angles of approach. Crystals are prisms that reflect externalities and refract within themselves, creating different colors, patterns, and arrays casting off in different directions. What we see depends on our angle of response—not triangulation but rather crystallization" (Richardson, in Richardson & St. Pierre, 2005, p. 963). However, from an interpretive-constructivist perspective, which is the basis of this book, triangulation remains a principal strategy to ensure validity and reliability.

A second common strategy for ensuring internal validity or credibility is *member checks.* Also called *respondent validation,* the idea here is that you solicit feedback on your preliminary or emerging *findings* from some of the people that you interviewed. "This is the single most important way of ruling out the possibility of misinterpreting the meaning of what participants say and do and the perspective they have on what is going on, as well as being an important way of identifying your own biases and misunderstanding of what you observed" (Maxwell, 2013, pp. 126–127). The process involved in member checks is to take your preliminary analysis back to some of the participants and ask whether your interpretation "rings true." Although you may have used different words (it is *your* interpretation, after all, but derived directly from their experience), participants should be able to recognize their experience in your interpretation or suggest some fine-tuning to better capture their perspectives. Some writers suggest doing member checks throughout the course of the study. Table 9.1 is a sample of the results from a member check. In this study, Crosby (2004) was interested in how learning experiences foster commitment to a career in teaching English as a foreign language. He asked several of his participants to comment on his findings regarding their experiences teaching English in a cross-cultural setting.

Adequate engagement in data collection is a third strategy that makes sense when you are trying to get as close as possible to participants' understanding of a phenomenon. How long one needs to observe or how many people need to be interviewed are always difficult questions to answer, since the answers are always dependent on the particular study itself. The best rule of thumb is that the data and emerging findings must feel saturated; that is, you

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

Table 9.1. Member Check Comments.

Name	*Comments*	*Action Taken*
Holly	"I think your statements are an accurate reflection of what I said and what my experience has been." The category you term "disorientating dilemma" puzzles me. That as a category doesn't quite ring true for me. Perhaps it came across that way, although I should also say that I'm not sure what you mean with that term and how it fits into learning experiences. Do you mean my challenges in teaching have encouraged/discouraged my commitment to teaching EFL?	Write back and explain about meaning of "disorientating dilemma" No action needed to change research results
Kate	"It was kind of fun to see a bunch of my own thoughts already categorized into a graphic!" Change spelling of Bombera to Bambara. Clarification of two phrases used as coding: Getting a Masters in TESOL, and looking for more teaching experiences.	Spelling corrected; phrases need not be adjusted
Grace	"I would agree with your categorization of comments." "I'd definitely agree with your conclusions." Charts gave "me greater insight into my own thinking."	No action needed
Mary	"Everything is right on! I have reviewed attachments and agree with what is written. The themes are accurate."	No action needed

(*continued*)

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

TABLE 9.1 (*Continued*)

Name	*Comments*	*Action Taken*
	"I really like the table; it was exciting to see my progression through your eyes."	
Ann	"I'd say it's pretty accurate. I can't think of anything I would add, change, etc."	No action needed
Shauna	"I do believe that the analysis rings true." "It was definitely an enlightening read. . . . It reminded me of certain convictions the Lord had placed on my heart to enter the field in the first place, and I feel encouraged as I look ahead towards my next step in the profession." "My commitment is first to God and His will for my life more so that [*sic*] my profession."	Note comment of commitment first to God then profession
Bob	"Both documents look great."	No action needed
Oliver	"When I left my interview with you I didn't feel like I expressed myself well, but after looking at your documents I think what you have is fine and rings true."	No action needed

Source: Crosby (2004). Reprinted with permission.

begin to see or hear the same things over and over again, and no new information surfaces as you collect more data.

Adequate time spent collecting data should also be coupled with purposefully looking for variation in the understanding of the phenomenon. Patton (2015) argues that credibility hinges partially on the integrity of the researcher, and one approach to dealing with this issue is for the researcher to "*look for data that support alternative explanations*" (p. 653, emphasis in original). He goes on to point out

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

that “failure to find strong supporting evidence for alternative ways of presenting the data or contrary explanations helps increase confidence in the initial, principal explanation you generated” (p. 654). Patton also reminds readers that there is often no clear-cut “yes” or “no” answer to whether data support an alternative explanation. Rather, “you’re searching for *the best fit,* the preponderance of evidence. This requires assessing the weight of evidence and looking for those patterns and conclusions that fit the preponderance of data” (p. 654, emphasis in original). Some writers even suggest that you should purposefully seek data that might disconfirm or challenge your expectations or emerging findings. This strategy has been labeled *negative* or *discrepant case analysis.*

Related to the integrity of the qualitative researcher is a fourth strategy sometimes labeled *researcher’s position,* or *reflexivity,* which is how the researcher affects and is affected by the research process (Probst & Berenson, 2014). Investigators need to explain their biases, dispositions, and assumptions regarding the research to be undertaken. Even in journal articles, authors are being called upon to articulate and clarify their assumptions, experiences, worldview, and theoretical orientation to the study at hand. Such a clarification allows the reader to better understand how the individual researcher might have arrived at the particular interpretation of the data. As Maxwell (2013, p. 124) explains, the reason for making your perspective, biases, and assumptions clear to the reader is not to eliminate “the researcher’s theories, beliefs, and perceptual lens. Instead, qualitative research is concerned with understanding how a *particular* researcher’s values and expectations influenced the conduct and conclusions of the study” (emphasis in original).

Yet another strategy is called *peer examination* or *peer review.* Certainly there’s a sense in which all graduate students have this process built into their thesis or dissertation committee, since each member of the committee reads and comments on the findings. A similar process takes place when an article is sent in to a peer-reviewed journal for publication; “peers” knowledgeable about the topic and the methodology review the manuscript and recommend publication (or do not). But such an examination or review can also be conducted by either a colleague familiar with the research or one new to the topic. There are advantages to both, but either way, a thorough peer examination would involve asking a colleague

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

to scan some of the raw data and assess whether the findings are plausible, based on the data.

Reliability or Consistency

Reliability refers to the extent to which research findings can be replicated. In other words, if the study is repeated, will it yield the same results? Reliability is problematic in the social sciences simply because human behavior is never static. Even those in the hard sciences are asking similar questions about the constancy of phenomena. Reliability in a research design is based on the assumption that there is a single reality and that studying it repeatedly will yield the same results. This is a central concept of traditional experimental research, which focuses on discovering causal relationships among variables and uncovering laws to explain phenomena.

Qualitative research, however, is not conducted so that the laws of human behavior can be isolated. Rather, researchers seek to describe and explain the world as those in the world experience it. Since there are many interpretations of what is happening, there is no benchmark by which to take repeated measures and establish reliability in the traditional sense. Wolcott (2005) underscores the inappropriateness of considering reliability in studying human behavior: "In order to achieve reliability in that technical sense, a researcher has to manipulate conditions so that replicability can be assessed. Ordinarily, fieldworkers do not try to make things happen at all, but whatever the circumstances, we most certainly cannot make them happen twice. And if something does happen more than once, we never for a minute insist that the repetition be exact" (p. 159).

Traditionally, reliability is the extent to which research findings can be replicated. In other words, if the study were repeated, would it yield the same results? Reliability is problematic in the social sciences simply because human behavior is never static, nor is what many experience necessarily more reliable than what one person experiences. All reports of personal experience are not necessarily unreliable, any more than all reports of events witnessed by a large number of people are reliable. Consider the magician who can fool the audience of hundreds but not the stagehand watching from the wings. Replication of a qualitative study will not yield the same

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

results, but this does not discredit the results of any particular study; there can be numerous interpretations of the same data. The more important question for qualitative research is *whether the results are consistent with the data collected.* Lincoln and Guba (1985) were the first to conceptualize reliability in qualitative research as "dependability" or "consistency." That is, rather than demanding that outsiders get the same results, a researcher wishes outsiders to concur that, given the data collected, the results make sense—they are consistent and dependable. The question then is not whether findings will be found again but whether the results are consistent with the data collected.

The connection between reliability and internal validity from a traditional perspective rests for some on the assumption that a study is more valid if repeated observations in the same study or replications of the entire study produce the same results. This logic relies on repetition for the establishment of truth, but as everyone knows, measurements, observation, and people can be repeatedly wrong. A thermometer may repeatedly record boiling water at 85 degrees Fahrenheit; it is very reliable, since the measurement is consistent, but not at all valid. And in the social sciences, simply because a number of people have experienced the same phenomenon does not make the observations more reliable.

It is interesting, however, that the notion of reliability with regard to instrumentation can be applied to qualitative research in a sense similar to its meaning in traditional research. Just as a quantitative researcher refines instruments and uses statistical techniques to ensure reliability, so too the human instrument can become more reliable through training and practice. Furthermore, the reliability of documents and personal accounts can be assessed through various techniques of analysis and triangulation.

Because what is being studied in the social world is assumed to be in flux, multifaceted, and highly contextual; because information gathered is a function of who gives it and how skilled the researcher is at getting it; and because the emergent design of a qualitative study precludes a priori controls, achieving reliability in the traditional sense is not only fanciful but impossible. Wolcott (2005) wonders whether we need "address reliability at all" other than to say why it is an inappropriate measure for assessing the rigor of a qualitative study. His objection is that "similarity of responses is taken to be the same as accuracy of responses," and we know that is a problematic assumption (p. 159).

Thus, for the reasons discussed, replication of a qualitative study will not yield the same results. As Tracy (2013) points out, "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" (p. 229). That fact, however, does not discredit the results of the original or subsequent studies. Several interpretations of the same data can be made, and all stand until directly contradicted by new evidence. So if the findings of a study are consistent with the data presented, the study can be considered dependable.

Strategies that a qualitative researcher can use to ensure consistency and dependability or reliability are triangulation, peer examination, investigator's position, and the audit trail. The first three have been discussed already under Internal Validity or Credibility. The use of multiple methods of collecting data (methods triangulation), for example, can be seen as a strategy for obtaining consistent and dependable data, as well as data that are most congruent with reality as understood by the participants. The audit trail is a method suggested by Lincoln and Guba (1985). Just as an auditor authenticates the accounts of a business, independent readers can authenticate the findings of a study by following the trail of the researcher. While "we cannot expect others to replicate our account," Dey (1993, p. 251) writes, "the best we can do is explain how we arrived at our results." Calling the audit trail a "log," as in what a captain might keep in detailing a ship's journey, Richards (2015) writes that "good qualitative research gets much of its claim to validity from the researcher's ability to show convincingly how they got there, and how they built confidence that this was the best account possible. This is why qualitative research has a special need for project history, in the form of a diary or log of processes" (p. 143).

An audit trail in a qualitative study describes in detail how data were collected, how categories were derived, and how decisions were made throughout the inquiry. In order to construct this trail, you as the researcher keep a research journal or records memos on the process of conducting the research as it is being undertaken. What exactly do you write in your journal or your memos? You write your reflections, your questions, and the decisions you make with regard to problems, issues, or ideas you encounter in collecting data. A running record of your interaction with the data as you engage in analysis and interpretation is also recommended. In a book-length or thesis-length report of the research, the audit trail is found in the methodology chapter (often with supporting appendixes). Essentially, it is a detailed account of how the study was conducted and how the data were analyzed. Due to space limitations, journal articles tend to have a very abbreviated audit trail or methodology section.

External Validity or Transferability

External validity is concerned with the extent to which the findings of one study can be applied to other situations. That is, how generalizable are the results of a research study? Guba and Lincoln (1981) point out that even to discuss the issue, the study must be internally valid, for "there is no point in asking whether meaningless information has any general applicability" (p. 115). Yet an investigator can go too far in controlling for factors that might influence outcomes, with the result that findings can be generalized only to other highly controlled, largely artificial situations.

The question of generalizability has plagued qualitative investigators for some time. Part of the difficulty lies in thinking of generalizability in the same way as do investigators using experimental or correlational designs. In these situations, the ability to generalize to other settings or people is ensured through a priori conditions such as assumptions of equivalency between the sample and population from which it was drawn, control of sample size, random sampling, and so on. Of course, even in these circumstances, generalizations are made within specified levels of confidence.

It has also been argued that applying generalizations from the aggregated data of enormous, random samples to individuals is hardly useful. A study might reveal, for example, that absenteeism is highly correlated with poor academic performance—that 80 percent of students with failing grades are found to be absent

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

more than half the time. If student Alice has been absent more than half the time, does it also mean that she is failing? There is no way to know without looking at her record. Actually, an individual case study of Alice would allow for a much better prediction of her academic performance, for then the particulars that are important to her situation could be discovered. The best that research from large random samples can do vis-à-vis an individual is to "make teachers and other clinicians more informed gamblers" (Donmoyer, 1990, p. 181). In qualitative research, a single case or a small, nonrandom, purposeful sample is selected precisely because the researcher wishes to understand the particular in depth, not to find out what is generally true of the many.

Although generalizability in the statistical sense (from a random sample to the population) cannot occur in qualitative research, that's not to say that nothing can be learned from a qualitative study. As Eisner (1998, pp. 103–104) points out, "generalization is a ubiquitous aspect" of our lives. However, "no one leads life by randomly selecting events in order to establish formal generalizations. We live and learn. We try to make sense out of the situations in and through which we live and to use what we learn to guide us in the future." As with internal validity and reliability, we need to think of generalizability in ways appropriate to the philosophical underpinnings of qualitative research.

Lincoln and Guba (1985) suggest the notion of *transferability*, in which "the burden of proof lies less with the original investigator than with the person seeking to make an application elsewhere. The original inquirer cannot know the sites to which transferability might be sought, but the appliers can and do." The investigator needs to provide "sufficient descriptive data" to make transferability possible (p. 298).

There are a number of understandings of generalizability that are more congruent with the worldview of qualitative research. Some argue that empirical generalizations are too lofty a goal for social science; instead, they say, we should think in terms of what Cronbach (1975) calls working hypotheses—hypotheses that reflect situation-specific conditions in a particular context. Working hypotheses that take account of local conditions can offer practitioners some guidance in making choices—the results of which can be monitored and evaluated in order to make better decisions in the future. Thus "when we give proper weight to local conditions, any generalization is a working hypothesis, not a conclusion" (p. 125). Patton (2015) also promotes the notion of extrapolating rather than making generalizations: "Unlike the usual meaning of the term *generalization,* an *extrapolation* clearly connotes that one has gone beyond the narrow confines of the data to *think about other applications of the findings.* Extrapolations are modest speculations on the likely applicability of findings to other situations under similar, but not identical, conditions. Extrapolations are logical, thoughtful, case derived and problem oriented rather than statistical and probabilistic" (p. 713, emphasis in original).

Modest extrapolations or working hypotheses are not the only way to think about generalizability in qualitative research. Erickson (1986) suggests the notion of "concrete universals" in which "the search is not for abstract universals arrived at by statistical generalizations from a sample to a population, but for concrete universals arrived at by studying a specific case in great detail and then comparing it with other cases studied in equally great detail" (p. 130). Every study, every case, every situation is theoretically an example of something else. The general lies in the particular; that is, what we learn in a particular situation we can transfer or generalize to similar situations subsequently encountered. This is, in fact, how most people cope with everyday life. You get one speeding ticket from a trooper pulling out from behind a billboard; subsequently, you slow down whenever you come upon a billboard on any road. You have taken a particular incident and formed a concrete universal. Erickson makes this same point with regard to teaching.

> When we see a particular instance of a teacher teaching, some aspects of what occurs are absolutely generic, that is, they apply cross-culturally and across human history to all teaching situations. This would be true despite tremendous variation in those situations—teaching that occurs outside school, teaching in other societies, teaching in which the teacher is much younger than the learners, teaching in Urdu, in Finnish, or in a mathematical language, teaching narrowly construed cognitive skills, or broadly construed social attitudes and beliefs.
>
> Each instance of a classroom is seen as its own unique system, which nonetheless displays universal properties of teaching. These properties are manifested in the concrete, however, not in the abstract. (p. 130)

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

The idea that the general resides in the particular, that we can extract a universal from a particular, is also what renders great literature and other art forms enduring. Although we may never live at the South Pole, we can understand loneliness by reading Byrd's account; and although we are not likely to be president, we can come up with concrete generalizations about power and corruption by listening to the Watergate tapes.

Probably the most common understanding of generalizability in qualitative research is to think in terms of the reader or user of the study. *Reader or user generalizability* involves leaving the extent to which a study's findings apply to other situations up to the people in those situations. The person who reads the study decides whether the findings can apply to his or her particular situation. This is a common practice in law and medicine, where the applicability of one case to another is determined by the practitioner. Nevertheless, the researcher has an obligation to provide enough detailed description of the study's context to enable readers to compare the "fit" with their situations.

Finally, Eisner (1998) argues that one of the stumbling blocks to our thinking about generalizability in the social sciences is the erroneous assumption that individual, nongeneralizable studies are limited in contributing to the accumulation of knowledge. However, knowledge is not inert material that "accumulates." Rather, he asserts, in qualitative research, accumulation is not vertical, but horizontal: "It is an expansion of our kit of conceptual tools" (p. 211). Connections between qualitative studies and one's world "have to be built by readers, who must . . . make generalizations by analogy and extrapolation, not by a watertight logic" (p. 211). "Human beings," Eisner writes, "have the spectacular capacity to go beyond the information given, to fill in gaps, to generate interpretations, to extrapolate, and to make inferences in order to construe meaning. Through this process knowledge is accumulated, perception refined, and meaning deepened" (p. 211).

To enhance the possibility of the results of a qualitative study "transferring" to another setting several strategies can be employed. The most commonly mentioned is the use of *rich, thick description.* Although *thick description,* "a phrase coined by the philosopher Gilbert Ryle (1949) and applied to ethnographic research by Geertz (1973)" originally meant an emic or insider's account (Maxwell, 2013, p. 138), it has come to be used to refer to a highly descriptive, detailed presentation of the setting and in particular, the findings of a study. Today, when rich, thick description is used as a strategy to enable transferability, it refers to a description of the setting and participants of the study, as well as a detailed description of the findings with adequate evidence presented in the form of quotes from participant interviews, field notes, and documents. As Lincoln and Guba (1985, p. 125) state, the best way to ensure the possibility of transferability is to create a "thick description of the sending context so that someone in a potential receiving context may assess the similarity between them and . . . the study."

Another strategy for enhancing transferability is to give careful attention to selecting the study sample. *Maximum variation* in the sample, whether it be the sites selected for a study or the participants interviewed, allows for the possibility of a greater range of application by readers or consumers of the research. As Patton (2015) notes, maximum variation sampling involves "purposefully picking a wide range of cases to get variation on dimensions of interest." There are two reasons for selecting a wide range of cases: "(1) to document diversity and (2) to identify important common patterns that are common across the diversity (cut through the noise of variation) on dimensions of interest" (p. 267). We would also add that including a variety of participants and/or sites in your study will enable more readers to apply your findings to their situation. Let's assume, for example, that you are a school principal interested in research on factors that promote community involvement in the school. The chances of your finding some helpful research are going to be increased if there's been a study that included a school in a community similar to yours. As another example, a qualitative study of the process and factors related to compliance with diabetes treatment will have more possibility of generalizing to more people if there was some variation in the characteristics of the participants (such as gender, age, education, length of time diagnosed).

Maximum variation is not the only sampling strategy one could use to enhance transferability. One could purposefully select a typical or modal sample. In typicality or modal category sampling, one describes how typical the program, event, or individual is compared with others in the same class, so that users can make comparisons with their own situations. In Wolcott's (2003) classic case study of an elementary school principal in the early 1970s, for example, he tells how he selected a principal who, "like the majority of elementary school principals" at the time of his study, would be male, responsible for one school, and "regard himself as a career principal" (p. 1).

Although maximum variation or typical sampling can be used to enhance transferability, there are certainly good reasons for studying a particular situation because of its uniqueness. And one would study the particular because there is something that can be learned from it, something that contributes, as Eisner (1998) noted in the quotes cited earlier, to the horizontal accumulation of knowledge. As Wolcott (2005, p. 167) points out, "every case is, in certain aspects, like all other cases, like some other cases, and like no other case."

Table 9.2 is a summary of the strategies discussed in this chapter for enhancing the rigor—indeed, the trustworthiness—of a qualitative study. These strategies are by no means inclusive of all that could be used, but they are some of the most commonly employed to ensure internal validity, reliability, and generalizability in interpretive qualitative research.

Most of the issues already described are appropriate considerations for validity and reliability in qualitative research designs in general. At the same time, some research designs require alternate and/or additional conceptualizations of validity in light of the purposes of the study. This is particularly the case for action research designs. As discussed in Chapter Three, the purpose of action research is to make something happen in order to solve a problem in practice. It is also to study the process of change itself. Hence, in addition to dealing with issues of validity and reliability in the ways described earlier, there are additional validity criteria particular to this form of research, including outcome validity, democratic validity, catalytic validity, and process validity (Herr & Anderson, 2015). Outcome validity is "the extent to which outcomes occur, which leads to a resolution of the problem that led to the study" (p. 67). Democratic validity refers to the extent to which the research is conducted in collaboration with the participants; catalytic validity refers to how the participants and researchers changed their views in the process. Process validity focuses on the

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

Table 9.2. Strategies for Promoting Validity and Reliability.

Strategy	*Description*
1. Triangulation	Using multiple investigators, sources of data, or data collection methods to confirm emerging findings.
2. Member checks/ Respondent validation	Taking tentative interpretations/findings back to the people from whom they were derived and asking if they are plausible.
3. Adequate engagement in data collection	Adequate time spent collecting data such that the data become "saturated"; this may involve seeking *discrepant* or *negative* cases.
4. Researcher's position or reflexivity	Critical self-reflection by the researcher regarding assumptions, worldview, biases, theoretical orientation, and relationship to the study that may affect the investigation.
5. Peer review/ examination	Discussions with colleagues regarding the process of study, the congruency of emerging findings with the raw data, and tentative interpretations.
6. Audit trail	A detailed account of the methods, procedures, and decision points in carrying out the study.
7. Rich, thick descriptions	Providing enough description to contextualize the study such that readers will be able to determine the extent to which their situations match the research context, and, hence, whether findings can be transferred.
8. Maximum variation	Purposefully seeking variation or diversity in sample selection to allow for a greater range of application of the findings by consumers of the research.

extent to which ongoing learning occurred during the *process* and stages of the research, as well as whether adequate evidence was provided to document the findings at each of those stages. While these additional criteria are important in action research,

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

essentially validity and reliability in any qualitative study are about providing information and rationale for the study's processes and adequate evidence so that readers can determine the results are trustworthy.

HOW ETHICAL CONSIDERATIONS RELATE TO THE TRUSTWORTHINESS OF QUALITATIVE RESEARCH

To a large extent, the validity and reliability of a study depend upon the ethics of the investigator. Patton (2015) identifies the credibility of the researcher along with rigorous methods as essential components to ensure the credibility of qualitative research: "ultimately, for better or worse, the trustworthiness of the data is tied directly to the trustworthiness of those who collect and analyze the data—and their demonstrated competence" (p. 706). It is the training, experience, and "intellectual rigor" of the researcher, then, that determines the credibility of a qualitative research study. "Methods do not ensure rigor. A research design does not ensure rigor. Analytical techniques and procedures do not ensure rigor. Rigor resides in, depends on, and is manifest in *rigorous* thinking—about everything, including methods and analysis" (p. 703). These qualities are essential because as in all research, we have to trust that the study was carried out with integrity and that it involves the ethical stance of the researcher. Suppose, for example, that you are studying an alternative high school reputed to have an unusually high student retention and graduation rate. You interview teachers, administrators, and students and begin to identify the factors that might account for the school's success. In reviewing some of the school records, you find that attendance and graduation rates have been inflated. Your decision as to how to handle this discovery will have a direct impact on the trustworthiness of your entire study. Although some sense of the researchers' values can be inferred from the statement of their assumptions and biases or from the audit trail, readers of course are likely never to know what ethical dilemmas were confronted and how they were dealt with. It is ultimately up to the individual researcher to proceed in as ethical a manner as possible.

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

Although policies, guidelines, and codes of ethics have been developed by the federal government, institutions, and professional associations, actual ethical practice comes down to the individual researcher's own values and ethics. Tracy (2013) suggests that ethical issues can exist with respect to procedures; that is, those guidelines "prescribed by certain organizational or institutional review boards (IRB) as being universal or necessary" (p. 243), such as "do no harm" and informed consent; they can be situational, such as those that come up in the research context; and they can be relational. "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" (p. 245). The protection of subjects from harm, the right to privacy, the notion of informed consent, and the issue of deception all need to be considered ahead of time, but once in the field, issues have to be resolved as they arise. This situational and relational nature of ethical dilemmas depends not upon a set of general preestablished guidelines but upon the investigator's own sensitivity and values.

In qualitative studies, ethical dilemmas are likely to emerge with regard to the collection of data and in the dissemination of findings. Overlaying both these processes is the researcher-participant relationship. For example, this relationship and the research purpose determine how much the researcher reveals about the actual purpose of the study—how informed the consent can actually be—and how much privacy and protection from harm is afforded the participants. Ethical considerations regarding the researcher's relationship to participants are a major source of discussion and debate in qualitative research, especially with the interest in critical, participatory, feminist, and postmodern research. When the research is highly collaborative, participatory, or political, ethical issues become prominent. Lincoln (1995) in particular aligns ethical considerations with the researcher's relationship with research participants and considers validity to be an ethical question. She suggests seven standards for validity, such as the extent to which the research allows all voices to be heard, the extent of reciprocity in the research relationship, and so on.

The standard data collection techniques of interviewing and of observation in qualitative research present their own ethical

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

dilemmas. As Stake (2005) observes, “Qualitative researchers are guests in the private spaces of the world. Their manners should be good and their code of ethics strict” (p. 459). Interviewing—whether it is highly structured with predetermined questions or semistructured and open-ended—carries with it both risks and benefits to the informants. Respondents may feel their privacy has been invaded, they may be embarrassed by certain questions, and they may tell things they had never intended to reveal.

In-depth interviewing may have unanticipated long-term effects. What are the residual effects of an interview with a teacher who articulates, for the first time perhaps, anger and frustration with his choice of career? Or the administrator who becomes aware of her own lack of career options through participation in a study of those options? Or the adult student who is asked to give reasons for failing to learn to read? Painful, debilitating memories may surface in an interview, even if the topic appears routine or benign.

However, an interview may improve the condition of respondents when, for example, they are asked to review their successes or are stimulated to act positively in their own behalf. Most people who agree to be interviewed enjoy sharing their knowledge, opinions, or experiences. Some gain valuable self-knowledge; for others the interview may be therapeutic—which brings up the issue of the researcher’s stance. Patton (2015) points out that the interviewer’s task “is first and foremost to gather data” (p. 495). The interviewer is neither a judge nor a therapist nor “a cold slab of granite—unresponsive to learning about great suffering and pain that may be reported and even re-experienced during an interview” (p. 495). Patton and others recommend being able to make referrals to resources for assistance in dealing with problems that may surface during an interview.

Observation, a second means of collecting data in a qualitative study, has its own ethical pitfalls, depending on the researcher’s involvement in the activity. Observations conducted without the awareness of those being observed raise ethical issues of privacy and informed consent. Webb, Campbell, Schwartz, and Sechrest (1981), in their book on nonreactive measures, suggest that there is a continuum of ethical issues based on how “public” the observed behavior is. At one end, and least susceptible to ethical violations, is the public behavior of public figures. At midposition are public

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

situations that "may be regarded as momentarily private," such as lovers in a park (p. 147). At the other end are situations involving "'spying' on private behavior," in which distinct ethical issues can be raised (p. 148).

Participant observation raises questions for both the researcher and those being studied. On the one hand, the act of observation itself may bring about changes in the activity, rendering it somewhat atypical. On the other, participants may become so accustomed to the researcher's presence that they may engage in activity they will later be embarrassed about, or reveal information they had not intended to disclose. Further, an observer may witness behavior that creates its own ethical dilemmas, especially behavior involving abuse or criminal activity. What if inappropriate physical contact between instructor and participant is witnessed while observing a volunteer CPR training session? Or a helpless teen is attacked by the group under study? Or a researcher witnesses utterly ineffective, perhaps potentially damaging counseling behavior? Knowing when and how to intervene is perhaps the most perplexing ethical dilemma facing qualitative investigators. Taylor and Bogdan (1984) conclude that although "the literature on research ethics generally supports a noninterventionist position in fieldwork," failure to act is itself "an ethical and political choice" (p. 71) that researchers must come to terms with.

Somewhat less problematic are the documents a researcher might use in a study. At least public records are open to anyone's scrutiny, and data are often in aggregated (and hence anonymous) form. But what of documents related to a continuing professional education program, for example, that reveal a misappropriation of funds? Or documents showing that administrative duties are based on certain favors being extended? And personal records pose potential problems unless they are willingly surrendered for research purposes.

Whether you are collecting data via interviews, observations, or documents, these sources of data in the *online* environment present additional ethical considerations such as how to obtain informed consent, assessing the authenticity of the data source, determining what is considered in the public domain and available to the researcher without consent, and so on. (See Chapter Seven for a fuller discussion of issues in online data collection.)

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

Analyzing data may present other ethical problems. Since the researcher is the primary instrument for data collection, data have been filtered through his or her particular theoretical position and biases. Deciding what is important—what should or should not be attended to when collecting and analyzing data—is almost always up to the investigator. Opportunities thus exist for excluding data contradictory to the investigator's views. Sometimes these biases are not readily apparent to the researcher. Nor are there practical guidelines for all the situations a researcher might face.

Disseminating findings can raise further ethical problems. If the research has been sponsored, the report is made to the sponsoring agency, and the investigator loses control over the data and its subsequent use. The question of anonymity is not particularly problematic in survey or experimental studies, when data are in aggregated form. At the other end of the continuum is a qualitative case study that, by definition, is an intensive investigation of a specific phenomenon of interest. The case may even have been selected because it was unique, unusual, or deviant in some way. At the local level, it is nearly impossible to protect the identity of either the case or the people involved. In addition, "The cloak of anonymity for characters may not work with insiders who can easily locate the individuals concerned or, what is even worse, claim that they can recognize them when they are, in fact, wrong" (Punch, 1994, p. 92).

This discussion on ethics in qualitative research has merely touched upon some of the issues that might arise when conducting this type of study. Readers interested in pursuing ethical considerations in more depth can turn to any number of sources. Patton (2015), for example, has a lengthy discussion and provides an "Ethical Issues Checklist" identifying the following 12 items to be considered when engaging in qualitative research:

1. Explaining the purpose of the inquiry and methods to be used
2. Reciprocity (what's in it for the interviewee and issues of compensation)
3. Promises
4. Risk assessment
5. Confidentiality
6. Informed consent

7. Data access and ownership
8. Interviewer mental health
9. Ethical advice (who will be your counselor on ethical matters)
10. Data collection boundaries
11. Ethical and methodological choices
12. Ethical versus legal (pp. 496–497)

In summary, part of ensuring for the trustworthiness of a study—its credibility—is that the researcher himself or herself is trustworthy in carrying out the study in as ethical a manner as possible.

Summary

As in any research, validity, reliability, and ethics are major concerns. Every researcher wants to contribute knowledge to the field that is believable and trustworthy. Since a qualitative approach to research is based upon different assumptions and a different worldview than traditional research, most writers argue for employing different criteria in assessing qualitative research.

The question of internal validity—the extent to which research findings are credible—is addressed by using triangulation, checking interpretations with individuals interviewed or observed, staying on site over a period of time, asking peers to comment on emerging findings, and clarifying researcher biases and assumptions. Reliability—the extent to which there is consistency in the findings—is enhanced by the investigator explaining the assumptions and theory underlying the study, by triangulating data, and by leaving an audit trail; that is, by describing in detail how the study was conducted and how the findings were derived from the data. Finally, the extent to which the findings of a qualitative study can be generalized or transferred to other situations—external validity—continues to be the object of much debate. Working hypotheses, concrete universals, and user or reader generalizability are discussed in this chapter as alternatives to the statistical notion of external validity. Rich, thick description facilitates transferability.

The trustworthiness of a qualitative study also depends on the credibility of the researcher. Although researchers can turn to guidelines and regulations for help in dealing with some of the ethical concerns likely to emerge in qualitative research, the

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use

burden of producing a study that has been conducted and disseminated in an ethical manner lies with the individual investigator.

No regulation can tell a researcher when the questioning of a respondent becomes coercive, when to intervene in abusive or illegal situations, or how to ensure that the study's findings will not be used to the detriment of those involved. The best a researcher can do is to be conscious of the ethical issues that pervade the research process and to examine his or her own philosophical orientation vis-à-vis these issues.

EBSCOhost - printed on 9/1/2024 1:10 PM via EDGEWOOD COLLEGE. All use subject to https://www.ebsco.com/terms-of-use