

2.3

Qualitative Rigor

How do I conduct qualitative research in a rigorous manner?

Williams, D. D. & Kimmons, R.

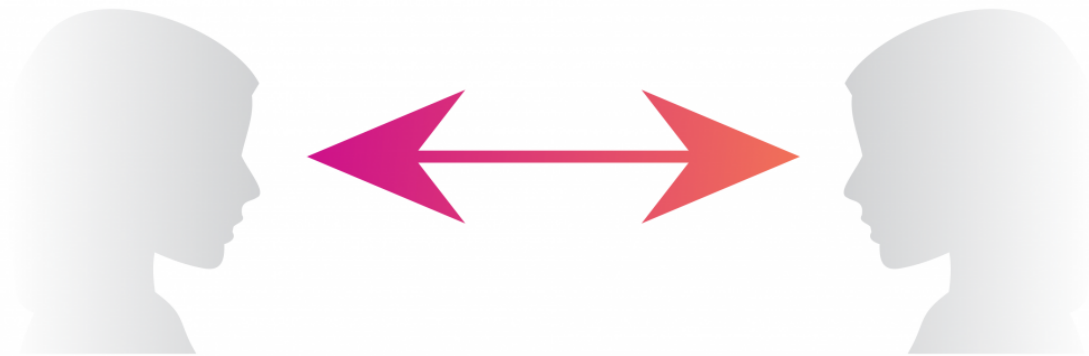
Qualitative Methods

Research Methods

Rigor

Trustworthiness

Qualitative research requires rigor to ensure that findings are believable and trustworthy. Four types of standards for ensuring trustworthiness are credibility, transferability, dependability, and confirmability. Credibility refers to the believability of study results to critical readers and participants. Techniques for enhancing credibility include prolonged engagement, persistent observation, triangulation, peer debriefing, negative case analysis, progressive subjectivity checks, and member checking. Prolonged engagement involves immersing oneself in the site or context of the study long enough to build trust with participants and overcome distortions due to one's presence. Persistent observation ensures depth of experience and understanding by exploring details of the phenomenon under study. Triangulation verifies findings through multiple sources of information, methods of data collection, and observations with multiple researchers. Peer debriefing involves meeting with a disinterested peer to question methods, emerging conclusions, and biases.



Imagine you are reading the abstract of a study, and the researcher claims something like "our interviews showed that games can be great for teaching collaboration skills."

What questions would you have for the researcher?

Would you believe them and trust them?

Or what would you need to know about how the study was conducted, who was interviewed, how the interviews were analyzed, and so forth before you would be willing to agree that "yes, indeed, games are great for teaching collaboration skills?"

Because qualitative inquiry is messy, relies on relatively few participant accounts, and often requires heavy interpretation on the part of the researcher, qualitative researchers need to be careful both in how they go about doing their work and how they report on it to ensure that their process is reasonable and their results are believable.

Key Terms

Confirmability

An expectation that results should be supported by participants, other researchers, and existing literature.

Credibility

An expectation that study results should be believable to critical readers, approved by participants, and otherwise true or accurate.

Dependability

An expectation that methods, logic, and reasoning guiding a study should be clear, stable, and consistent.

Discipline

An expectation that the researcher is thoughtful and methodical, following key standards and norms that are generally accepted by other researchers who use similar methodologies.

Generalizability

The ability to take the results of a study focusing on a sample of a population and to apply them to the overall population.

Rigor

An expectation (common in all research methodologies) that the researcher is being thorough, responsible, reasonable, and accurate.

Transferability

The ability of a reader to apply or transfer the results of a study to their own situation or context.

Trustworthiness

An expectation in qualitative methodologies that the researcher should provide enough explanation, transparency, and evidence that their results can be confidently believed.

Earlier we established that research is both systematic and auditable. Embedded in this is the notion that our practices and the results of our efforts should be rigorous, meaning that we strictly adhere to specific criteria of quality. In qualitative methodologies, this rigor typically means two things: (a) that we are disciplined ([Cronbach & Suppes, 1969](#)), meaning that we follow certain key standards established by other researchers, and (b) that we are trustworthy ([Guba & Lincoln, 1989](#); [Lincoln & Guba, 1985](#)), meaning that whoever reads our work should be able to feel informed and confident in both what we did and what we concluded. For these reasons, qualitative researchers have developed a variety of standards to help ensure both discipline and trustworthiness, which we will proceed in this chapter to explain in more detail.

The rationale for trustworthiness as the central objective of these standards is centered on the desire most people have for truth. Qualitative researchers agree that most claims people make are based on their subjective constructions of reality. A major objective in sharing our findings from inquiry thus becomes the persuasion of others that our constructions of reality are of value and should also be considered in their constructions. Whether or not these claims are true in any ultimate sense can only be tested over time through many different experiences in a variety of contexts (cf., intersubjectivity), but for any given study, the objective is one of persuasiveness — providing evidence that is compelling enough that audiences are willing to listen to and consider the claims made. In other words, the more the researcher can do to make the inquiry trustworthy, the more likely it is that readers will be persuaded to read on.

The standards presented by Lincoln and Guba (1985) and by Guba and Lincoln (1989) provide an excellent core of standards for a beginning. They suggested four types of standards be used to ensure trustworthiness: credibility, transferability, dependability, and confirmability. They also recommended several techniques for conducting studies so that they meet these standards. Although no single study is likely to adhere to all of these standards, which we will discuss below, as we meet or address more standards we make our work more believable and increasingly influential to people who will read it.

Learning Check

Which of the following are rigor requirements in qualitative research?

- ☐ Validity
- ☐ Confirmability
- ☐ Trustworthiness
- ☐ Dependability

Credibility

Credibility is the standard by which a qualitative study is expected to be believable to critical readers and to be approved by the persons who provided the information gathered during the study. Lincoln and Guba recommended several techniques researchers may use to enhance the credibility of their research, including prolonged engagement, persistent observation, triangulation, peer debriefing, negative case analysis, progressive subjectivity checks, and member checking, which we will now explain.

Prolonged Engagement

Prolonged Engagement is a technique whereby researchers immerse themselves in the site or context of the study long enough to build trust with the participants and for the researcher to experience the breadth of variation and to overcome distortions due to their presence (cf., Hawthorne Effect). This may mean an entire year or longer for some studies or as little as a month for others, depending on the size of the study and the level of depth needed for the researcher to become part of a

community and understand what is happening. There is no set amount of time a qualitative inquiry should last, but the proper length can be estimated by the researcher once they have spent some time in the site.

For example, if a researcher wanted to understand the phenomenon of Texas high school football, this would require being present at least through a full season of the sport and may also require presence in the pre-season and the off-season, whereas a researcher who simply showed up for a championship game would have little understanding of the nuances, histories, difficulties, perplexities, and larger context of what they were witnessing. If a researcher can be present in a setting long enough to see the range of things to be expected in such a site (e.g., not *just* the championship game), then the results produced will be more credible.

Persistent Observation

Persistent Observation is a technique that ensures depth of experience and understanding in addition to the broad scope encouraged through prolonged engagement. To be persistent, the researcher must explore details of the phenomenon under study to a deep enough level that they can decide what is important and what is irrelevant and focus on the most relevant aspects.

For instance, if a researcher wanted to understand the impacts of homework on marginalized students, they might begin by talking to students at school but would likely need to observe them after school or at home as well. Issues that such students might face, such as working jobs to support their families, providing childcare to working parents, not having access to a quiet study location, and so forth, would not be readily apparent if the researcher didn't leave the confines of the school or the 9 am to 3 pm hours of the established school day.

Failure to engage in persistent observation would mean that the researcher would learn very little detail about particular aspects of the phenomenon under study. Even if the researcher engaged in prolonged engagement (i.e., sufficient time), this alone would not mean that they had explored the phenomenon in sufficient depth, persistently learning more about the phenomenon from a variety of angles and in a variety of ways. Without such persistence, results would be limited in scope and less credible.

Triangulation

Triangulation is the verification of findings through (a) referring to multiple sources of information (including literature), (b) using multiple methods of data collection, and often (c) conducting observations with multiple researchers. If a conclusion is based on one person's report, given during one interview to only one interviewer, then it will be less credible than if several people confirmed the finding at different points in time, during multiple interviews, through various unstructured observations, in response to queries from several independent researchers, and in the review of literature.

For instance, if a researcher wanted to understand racially-motivated bullying in schools, a single interview with a bullied child might provide rich data about that child's experience, but readers might be left wondering how pervasive such experiences are and how race played a role in the child being targeted. If, however, the researcher can interview multiple students who have had similar experiences, connect this to other datasets (such as instances of bullying in school behavioral referrals), connect this to previous studies or theoretical literature, and find similar results from other researchers in other contexts, then the results will be more credible to a reader. Although all three forms of triangulation are not required for every conclusion the researcher makes, credibility is increased as more and more triangulation occurs.

Peer Debriefing

Peer Debriefing is a technique whereby a researcher meets with a disinterested peer so that the peer can question the researcher's methods, emerging conclusions, and biases. A disinterested peer might include anyone who is willing to ask

probing questions and who is not a participant or researcher in the setting where the study is being conducted.

For instance, a researcher studying the experiences of immigrant students in rural communities might periodically meet with a colleague who studies early childhood literacy. This peer would then challenge the assumptions that the researcher is starting to make about the context or results and would encourage them to both think critically about what they are observing and to figure out how to make results understandable to someone who is outside the context.

This technique is meant to keep the researcher honest by having someone else independently point out the implications of what they are doing. If a researcher can provide evidence of having engaged in peer debriefing and can show the reader how the report was modified through the influence of the peer, credibility is improved.

Negative Case Analysis

Negative Case Analysis is an analytic procedure that is meant to refine conclusions until they account for all possible cases, without any exceptions. The process involves developing hypotheses based on extensive fieldwork and then searching for cases or instances within the site under study that may contradict the conclusions proposed by the hypotheses.

For instance, if a researcher begins to conclude that poverty is having a serious, negative impact on student achievement in a community, then they should seek to find any negative cases in which a child from an impoverished family is actually excelling academically. This will help them to better understand the interaction between poverty and achievement and what some individuals or families must do to mitigate it.

If no contradictory cases are found after extensive searching, then the hypotheses are considered more credible because no evidence has been found to negate them. If such evidence *is* found, however, then the hypotheses are modified to account for the new data associated with the negative cases. This process continues until the hypotheses have been modified to account for all negative cases and no new negative cases can be found.

If a researcher completes such an extensive process, the resulting qualitative inquiry report is considered very credible. Though single studies sometimes fail to account for negative cases due to limited exposure or familiarity with the topic, as researchers engage in series of studies over time on the same topic, it is expected that they will eventually grapple with negative cases in a robust way.

Progressive Subjectivity Checking

Progressive Subjectivity Checking is a technique whereby a researcher archives their changing expectations and assumptions for a study. These expectations might include *a priori* and emerging constructions or interpretations of what is being learned or what is going on.

For instance, if a female academic is studying gender disparities in higher education, then she might bring to light her own experiences of inequity, memo how her participants' experiences reflect, reinforce, or challenge her own assumptions and experiences, and acknowledge how her thinking on the topic is changing as she progresses through the study.

In all qualitative research, the researcher is responsible for revealing their biases and preferences in reports, field notes, and the audit trail both initially and over time, and as Guba and Lincoln (1989) explain "if the [researcher] 'finds' only what he or she expected to find, initially, or seems to become 'stuck' or 'frozen' on some intermediate construction [interpretation], credibility suffers" (p. 238).

Emic Perspective

Emic Perspective or the folk perspective of participants is the insider view of how participants see and understand themselves from the inside-out, and researchers improve credibility by showing that they are able to understand and communicate about the phenomenon being studied as an insider.

For instance, any researcher studying young children would not just need to be able to interpret behaviors through etic theoretical or psychological lenses (e.g., phobias, apathy, intelligence) but would also need to show that they understand the children as they understand themselves, often using their same words (e.g., "love," "hate," "smart," "dumb," "mad").

That is, it should be clear to readers that the researcher discovered something of the viewpoints held by the people they studied and can see them as they see themselves. If *only* the researcher's outsider perspective is present, then the study will lack one of the most critical characteristics of a qualitative study: the type of understanding that can only come from empathy.

This also helps ensure that the researcher is not pigeonholed by *a priori* assumptions that might blind them to new discoveries. If the researcher's original hypotheses are simply confirmed, then qualitative inquiry probably is not the appropriate approach to use, but by discovering emic perspectives, researchers can add richness to existing understandings of phenomena and make their results more credible.

Learning Check

Which of the following would be an example of researchers employing emic perspectives?

- ☐ Observing a kindergarten classroom through a one-way mirror
- ☐ Explaining teenagers' moral reasoning in their own words
- ☐ Utilizing a standardized test to determine intelligence
- ☐ Analyzing usage data for a web-based learning app

Member Checking

Member Checking is a technique whereby a researcher provides the data record, interpretations, and/or reports for review by the participants who provided the data - the natives. This validates that the represented emic perspective is accurate and is one of the most important techniques for ensuring credibility.

For instance, if a researcher interviewed parents about their reasons for enrolling their students in charter schools, the researcher might provide each parent with a transcript of the interview as well as the researcher's summary and key takeaways of what was said.

This allows the participant to either tell the researcher "Yes, you got it!" or "No, I actually meant something else." If they agree that their perspectives have been adequately represented and that the conclusions reached in the report are accurate to them, then the reader will be more convinced that the qualitative inquiry itself is credible.

However, because member checking might require participants to read, understand, and provide feedback on data and results, researchers may need to employ some creativity and empathy in how they go about doing member checks with diverse participants. Young children might not be able to read, second language learners may have difficulty understanding reports, and the lay public may not understand technical terms. In such situations, the researcher might find alternative ways to share what they are concluding in understandable ways, perhaps simplifying reports or reading segments to a participant and then relying on oral feedback and reactions.

Transferability

Because qualitative studies are not designed to be generalizable like quantitative studies, their results should never be framed as universal truths or as conclusions that are true in all contexts and settings. Attempting to generalize is a common mistake in qualitative research and can signal to your readers that you are not aware of the limits of your own methods and do not recognize the actual complexity of the phenomena you are studying. To avoid generalization, be sure that you are using softening words, like "may," "suggests," "perhaps," and so forth, and that you are aware that what you see in your research setting may not always be true in other research settings (or even in the same setting at a different time).

Yet, given sufficient detail, qualitative studies can provide insight into what is happening in new contexts that you, as the researcher, may not be aware of. **Transferability** is the standard by which qualitative study results are expected to be able to be transferred or applied to new, novel contexts.

In other words, the qualitative researcher should consider whether their findings, which were discovered in one situated context, can apply to other contexts or settings as well (such as where the reader is working). Whether findings can be transferred or not is an empirical question, which cannot be answered by the researcher alone, because the reader's context must be compared to the research context to identify similarities. The more similar, the more likely it is that the findings will be transferable. Thus, readers must be the ones to determine whether the qualitative inquiry is transferable, not the researcher.

The researcher is expected to facilitate transferability, however, by providing clear descriptions of the time and context in which results and conclusions are developed, providing thick descriptions of the phenomena under study, and providing as much explanation about the context in which the study took place as possible. In short, more details give readers more power to discern which results might transfer to their contexts and which might not, and the rigorous qualitative researcher provides readers with sufficient detail to determine for themselves whether study results will transfer to their unique contexts.

Learning Check

How is transferability different from generalizability?

- ☐ They are synonyms or have the same meaning.
- ☐ Transferability is contextual, whereas generalizability is universal (to a population).
- ☐ Transferability requires significance testing, whereas generalizability requires proper sampling.

- ☐ Transferability is the first step toward generalizability and can become generalizability if done properly and often enough.

Dependability

Dependability is the standard by which the logic, reasoning, methods, and results are expected to be stable or consistent over time. To check the dependability of a qualitative study, one looks to see if the researcher has been careless or made mistakes in conceptualizing the study, collecting the data, interpreting the findings and reporting results. The logic used for selecting people and events to observe, interview, and include in the study should be clearly presented. The more consistent the researcher has been in this research process, the more dependable are the results.

For instance, a study that was attempting to understand African American students' experiences in an inner-city school but then shifted to interviewing white students, rural students, etc. would have deviated from the established reasoning and methods proposed in the study. Such deviations often occur out of convenience to the researcher (e.g., a target population is no longer available for study), but they represent a serious threat to dependability.

A major technique for assessing dependability is a dependability audit in which an independent auditor reviews the activities of the researcher (as recorded in an audit trail in field notes, archives, and reports) to see how well the techniques for meeting the credibility and transferability standards have been followed. If the researcher does not maintain any kind of audit trail, then the dependability cannot be assessed, thereby diminishing it along with overall trustworthiness.

Confirmability

Confirmability is the standard by which a qualitative study is expected to be supported by informants (participants) who are involved in the study and by events that are independent of the researcher. Reference to literature and findings by other authors that confirm the researcher's interpretations can strengthen confirmability of the study in addition to information and interpretations by people other than the researcher from within the inquiry site itself.

For instance, if a researcher studied a few women's experiences in computer science and found that they felt empowered and treated equally to their male counterparts, but many external examples of harassment and mistreatment of women were arising both in the literature and mainstream news, then this would lead the reader to wonder whether the researcher's findings actually represented the real experiences of women in computer science or were merely an anomaly or a misinterpretation. In this case, it seems possible that a researcher was misinterpreting the experiences of women they were interviewing or simply didn't interview enough women to understand the issue fully.

This does not mean that qualitative research results must always agree with all other sources of information, but it does mean that there should be ways to confirm research results, either by reviewing data sources (such as transcripts), repeating the study in different contexts, or comparing results to other evidence.

To do this, a confirmability audit can be conducted at the same time as the dependability audit, as the auditor asks if the data and interpretations made by the researcher are supported by material in the audit trail, are internally coherent, and represent

more than "figments of the researcher's imagination" (Guba & Lincoln, 1989, p. 243). If such an audit attests to the confirmability of the study, it is more likely to be accepted by readers.

Other Criteria

In addition to the standards discussed above, several other important considerations are suggested in the literature, including meaningfulness, appropriateness, natural conditions, ethical treatment, and audit trails.

Meaningfulness

Meaningfulness is the expectation that a study will address a worthwhile problem or issue, and if it doesn't, then it is not worth doing. This holds true for all research, not just qualitative inquiry. There should be a rationale providing justification for the time, money, and other resources devoted to the study. Deciding whether a problem is meaningful or not is a subjective determination, but the researcher can provide evidence and logic to support his or her decision, which will allow the reader to make an informed decision as to whether the study merits attention.

Appropriateness

Appropriateness is the expectation that a study's methods align with its intended goals. Not all research is or should be qualitative. If the needs call for it and the researcher can justify the application of a qualitative approach, then qualitative methods can be reasonably used. If the goal or need is something else, such as generalizability or design of an intervention, then qualitative methods alone are likely not appropriate.

The danger here is that researchers might approach qualitative research with a means-oriented mindset, wherein they apply qualitative methodologies inappropriately to problems that they are not equipped to solve, thereby overstepping the limits of the paradigm and failing the incommensurability test.

Studies should, therefore, provide a rationale to readers that both the goal of the research study is meaningful (i.e., meaningful ends) and also that qualitative methods are the right or best way of achieving the intended goal (i.e., appropriate means).

Natural Conditions

The expectation of **natural conditions** means that studies should be conducted under the most natural conditions possible. Manipulation of the participants through random assignment, submission to unnatural measurement instruments, or exposure to unnatural treatments should be avoided. The researcher should be as unobtrusive as possible so participants are acting essentially as they would if the researcher were simply another participant in the setting and not also conducting inquiry.

Ethical Treatment

Though all research should follow **ethics** guidelines, qualitative research places an especially high emphasis on valuing participant self-determination and social and psychological wellbeing. This may mean that practices sometimes used in other research projects, such as deception, may not be appropriate in qualitative settings and indeed may reduce the legitimacy of the qualitative approach.

This generally means that participants should be given the opportunity to react to the data record and have their disagreements with the researcher's interpretations taken seriously. Participants should also be given anonymity in any reports, and there should be no indications that participants were treated with disrespect or cruelty.

Audit Trail

An **audit trail** is simply the records kept of how a qualitative study was conducted. The audit trail should include all field notes and any other records kept of what the researcher does, sees, hears, thinks, etc. These notes describe the researcher's evolving relationship to what they are observing and what is being learned, and they also describe the researcher's thoughts about how to proceed with the study, sampling decisions, ethical concerns, and so on.

Each researcher is free to create a unique audit trail that fits the study being conducted, and the audit trail may be used as a reference throughout the study to review what has been done and to consider alternative plans, in addition to serving as part of the dependability and confirmability audits described above. Often, audit trails and field notes are the same, and if field notes are kept current and are easily accessible, no extra audit trail may be necessary (although some people like to keep a separate file for audit trail documentation).

To help an auditor, many researchers create a brief chronological index to their study. They list choices they made each day of the study, actions they engaged in, and some of their thoughts about how the study is going at each stage. The auditor can then go from this listing to the field notes, audio and video recordings, and other files associated with the inquiry to reconstruct how the study was conducted, to understand how conclusions were reached, and to make the dependability and confirmability judgments described earlier.

Signaling Rigor

In addition to following these standards, **rigor** must also be effectively communicated to the reader in order to serve its purpose. For this reason, reports should be well written to include description, analysis, and synthesis (cf., [Wolcott, 1994](#)), as well as to reveal the biases and assumptions of the researchers involved. Attempts to share what the researcher is learning should be communicated clearly. Descriptions should develop a sense of "being there" for the reader. Analyses should be logically presented. The audience for the report should be identified, and the report should address the concerns of that audience. The grammar and use of language should be of the highest quality.

Although necessary balances between description, analysis, and synthesis will vary depending on the length of the report and the purposes of the inquiry, readers need to have some raw description of scenes from the research site to use in judging the conclusions that are reached and to make their own conclusions independently. They also should see some synthesis of results by the researcher, in which all contradictions in findings are analyzed and/or resolved. Although there are paradoxes in the world, a report that presents conflicting pieces of evidence without discussing them and trying to discern their nature (i.e., whether it is a true paradox or whether one side of the issue is erroneous) needs to be improved.

Relevant characteristics of the researcher should also be clearly revealed so that the reader can understand the context from which the study emerged more completely. This may be done either *explicitly* in an appendix, in the foreword, or in the body of the text. Or it may be done *implicitly* in the text as the researcher describes his or her methods, decisions, reasons for doing the study, and so on.

Also, as researchers employ any of the techniques described above, they should explicitly state so in the methods sections of their reports and only mention techniques that they intentionally employed. This should consist of more than a list but should include necessary descriptions, such as how member checking was conducted, who the peer debriefers were, what some examples of negative cases were, and so forth. By doing this, qualitative researchers can better establish the rigor of their work to discerning and critical readers and also legitimize their processes and results as being worthy of consideration.

References

Cronbach, L. J., & Suppes, P. (1969). *Research for tomorrow's schools*. New York: Macmillan.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Lincoln, Y. S., & Guba, E. G. (1985). *Qualitative inquiry*. Beverly Hills, CA: Sage.

Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Sage.



This work is released under a CC BY license, which means that you are free to do with it as you please as long as you properly attribute it.

This content was previously published as follows:

Williams, D. D. (2018). *Qualitative Inquiry in Daily Life* (1st ed.). EdTech Books.
Retrieved from <https://edtechbooks.org/qualitativeinquiry>

