The Quest for Rigor in Qualitative Studies: Strategies for Institutional Researchers

By Felice Billups, Ed.D., Professor in the Educational Leadership Doctoral Program, Johnson & Wales University

Author’s note: An earlier version of this article was printed in The NERA Researcher in 2014.

Institutional researchers devote considerable effort to ensure that their studies are rigorous, valid, reliable, and actionable. ‘Clients’ (e.g. faculty, staff) expect these studies to be professionally accomplished with precision and objectivity, grounded in sound ethical practice. These expectations, however, are typically oriented towards quantitative designs; the concepts of rigor, validity, reliability, and generalizability are not applied to qualitative studies in the same ways. If different standards are applied to qualitative research to ensure quality and rigor, how can institutional researchers apply those standards to evaluate qualitative studies?

In this brief overview I will outline the definitions and descriptions of trustworthiness, followed by recommendations for ways in which institutional researchers can incorporate these strategies into their qualitative studies.

What is Trustworthiness?

Trustworthiness, a concept adapted and promoted by Lincoln and Guba (1985), is considered the quintessential framework for evaluating qualitative research, but receives minimal attention from many institutional researchers, especially those predominantly oriented to quantitative methods. In fact, many researchers expect that the same principles of validity, reliability, and generalizability can and should be applied to qualitative designs. This is not typically the case, and understanding the evaluative criteria for assessing qualitative research is a necessary component in any institutional researcher’s toolkit.

Four elements comprise the original trustworthiness framework: credibility (truth), dependability (consistency), transferability (applicability), and confirmability (neutrality). Authenticity, a fifth element added since the original discussion, is endorsed by some qualitative researchers as an equally important evaluative element (Polit & Beck, 2011). These elements are explained below, preceded by a chart that illustrates the association of these elements compared with the quantitative paradigm:

<table>
<thead>
<tr>
<th>Quantitative</th>
<th>Qualitative</th>
<th>‘Translation’</th>
<th>Essential Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity</td>
<td>Credibility</td>
<td>Truth</td>
<td>Are results believable? Do they appear truthful?</td>
</tr>
<tr>
<td>Reliability</td>
<td>Dependability</td>
<td>Replicable</td>
<td>Are results consistent over time?</td>
</tr>
<tr>
<td>Generalizability</td>
<td>Transferability</td>
<td>Applicable</td>
<td>Are the results applicable to similar settings?</td>
</tr>
<tr>
<td>Objectivity</td>
<td>Confirmability</td>
<td>Neutrality</td>
<td>Corroborated through triangulation?</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Authenticity</td>
<td>Reality(ies)</td>
<td>Are all ‘realities’ represented?</td>
</tr>
</tbody>
</table>

While this chart illustrates the progression from the quantitative to qualitative frameworks for evaluative criteria, the terms cannot be matched exactly; concepts are broadly transferable within the research paradigms.

Trustworthiness in Practice

The concept of rigor in qualitative research studies has been debated for many decades. For institutional researchers, applying trustworthiness strategies not only demonstrates sound practice but also adds the elements of outreach, engagement, and collaboration. The practice of working closely with participants, colleagues, and fellow researchers adds depth to a study less achievable with other approaches. To assist practitioners, each element of trustworthiness is described below, followed by examples of how these concepts can be applied to institutional research projects.

Credibility: Are the qualitative findings believable? Do the findings appear truthful, capturing a holistic representation of the phenomenon under exploration?
**Prolonged engagement**, in conjunction with **persistent observation** (intense focus on the aspects of setting and phenomenon), suggests that the researcher must spend considerable time in the field to thoroughly understand participant perspectives and to offset the researcher’s own bias (although how much time depends on the nature of the study (Wallendorf & Belk, 1989)). As Lincoln and Guba (1985) note, “if prolonged engagement provides scope, persistent observation provides depth” (p. 304).

**IR Applications**: The concept of prolonged engagement does not translate into conducting a particularly long interview; instead, the researcher needs to develop a close familiarity with participants as they live or work in their ‘natural’ setting (classroom, residence hall, campus locales). When conducting a study, institutional researchers might consider establishing a series of gatherings (luncheons, brainstorming sessions) before conducting interviews or focus groups, to engage the target population and to better understand their behaviors and perspectives through observation and time spent together.

**Peer debriefing** involves securing feedback from another researcher to compare conclusions; peers may address questions of bias, errors of fact, competing interpretations, convergence between data and phenomena, and the emergence of themes, all of which comprise a lengthy but important process for reinforcing credibility.

**IR Applications**: Working with other IR staff or partnering with faculty to review your research, particularly involving those individuals for whom or with whom the research will benefit, is an excellent way to strengthen verity. This strategy includes the added value of developing collaborative partnerships, or encouraging other campus offices to include the IR office in their work. I recommend using individuals from within the campus community for this practice, as they are most likely the individuals who know the target population and can most easily identify contradictions or errors in fact. There may also be instances where you can ask a colleague to review your conclusions or executive summary. Be judicious and apply this strategy sparingly!

**Member-checking**, different from peer debriefing, is when select participants are asked to review the findings or preliminary analysis to assess whether those findings reflect what they expressed to the researcher. This feedback may be obtained either in writing or in face-to-face conversations, and is also a way to gather additional data. Member checking is, however, considered controversial for many reasons and is a process which must be carefully applied (Lincoln & Guba, 1985; Polit & Beck, 2011) as some experts hint that the practice may corrupt the data.

**IR Applications**: One approach may be to identify a few carefully chosen participants who you believe will provide the best feedback, and ask those individuals to comment on whether your summary truly reflects the essence of the interview or focus group session. By providing a few questions to frame their response, and by providing a date by which you would like to hear back from them, you create an opportunity for participants to reinforce and ‘own’ the findings. This feedback loop may work best via email when involving students, faculty, or staff; it makes good use of time and technology to be as efficient as possible when conducting member checking with campus groups. In the end, it is not necessary to send all of your raw data, or all of your transcriptions, or all of your findings, for member corroboration – just share portions of your results that can be easily digested and reviewed by participants who will provide honest feedback.

**Triangulation** involves using multiple data sources to produce greater depth and breadth of understanding. This effort corroborates findings and/or builds a more holistic picture of the phenomenon. Triangulation is accomplished through one or more of the following approaches: methods triangulation (different data collection methods including interviews, journals, focus groups, observations, and documents), data triangulation (using different participants or data sources within one study in subsets of people, time, and space)(Cohen & Crabtree, 2006), researcher/analyst triangulation (using multiple analysts to review findings through cross-case or within case analyses), or theory triangulation (using multiple perspectives to interpret the data).

**IR Applications**: Institutional researchers may feel that they do not have time to collect data from more than 1-2 sources; triangulation is a wonderful goal but not always achievable. This strategy may be easier to apply than it appears, though, and may be easily realized. For instance, conducting a few elite interviews with key people at the conclusion of a survey questionnaire administration is one strategy; this would represent methods triangulation.
Studying one particular population of students about their satisfaction with campus services could be triangulated by administering a survey questionnaire, conducting focus groups, and asking for reflective questions to be submitted, all from the same population, constituting data triangulation. Theory or analyst triangulation may be more time-consuming for an institutional research project but you could involve a fellow IR staff member by asking them to serve as a triangulation analyst, doubling also as a peer debriefer (see above). These activities takes less time than it appears and the benefits allow for a richer and deeper understanding of the results.

**Negative case analyses** apply disconfirming evidence to search for other interpretations in a study; the use of conflicting findings builds a richer picture of the phenomenon and allows for continuous refinement of the results (Creswell, 2013).

**IR Applications**: Applying this strategy may be as easy as selecting a few participants who you know will provide opposing viewpoints and involve them in your sample. Whenever you conduct research with students, faculty, or even staff members, there are always voices of discontent and their perspective should be included in the final results. I once conducted a study with faculty regarding their opinions about a campus building project, yet the only ones who attended the focus group were the ones in favor of the project. Faculty who strongly opposed the project declined to participate because of their fear of publicly stating their opinions. Promising confidentiality, I approached some of these individuals and asked them about their reasons for disliking the project. The end result is that I uncovered nuances about both the positive opinions as well as the negative opinions that made my results more actionable.

**Dependability**: Are the findings stable and consistent over time and across conditions? Would the same data collection methods yield the same or similar results? Ensuring that the same research process generates the same essential findings often depends on external audits, these audits (also known as inquiry audits) are an important strategy for feedback, to assess the truthfulness of preliminary findings (Miles & Huberman, 2014). The drawbacks to this strategy, however, may produce conflict between two researchers’ perspectives, in which case the primary researcher may decide to revise the study.

**IR Applications**: There are several options for external auditing: 1) you can ask institutional research colleagues at other institutions to review all of your research procedures and findings, 2) you can ask a faculty member or team of faculty to collaborate, building partnerships to ensure subsequent acceptance of the findings, or 3) you can engage graduate students with research experience to serve as the external auditors. In any of these instances, you generate alternate perspectives on your study to develop ‘truthfulness’ and develop partnerships.

**Transferability**: Are these findings comparable? While the goal of qualitative research is not to produce results which are statistically generalizable, the intent is to produce findings which other researchers can interpret for similar settings, even to the point of applying the research design for their own purposes (Trochim, 2006). As Lincoln and Guba note (1985), “by describing a phenomenon in sufficient detail, one can begin to evaluate the extent to which the conclusions drawn are transferable to other times, settings, situations, and people” (p.306).

This work is accomplished through the strategy of ‘thick description’, a term first used by Ryle (1949) and extended by Geertz (1973). ‘Thick description’ comprises the researcher’s field notes including extensive detail and explicit descriptions when recording conversations, observations, and interpretations during data collection. This practice allows the researcher to evaluate comparable transferability of the same circumstance of people, place, and phenomenon under similar conditions, with similar participants.

**IR Applications**: Institutional researchers often fall into the trap of letting efficiency and expediency supersede comprehensiveness. I have often been tempted to limit an interview to 60 minutes, when the conversation might have yielded greater detail or richer descriptions of the participant’s experience regarding the phenomena if I had extended the session. The careful construction and use of qualitative instrumentation, supported by probes or prompts for observation/field note rubrics, is essential to ensuring transferability. The more detailed and nuanced information you generate, the greater the likelihood that your findings can be applied to a similar setting, population, or case. The concept of generalizability is not applied here as it would be in a quantitative study, but the relevance of transferring the findings is no less important.
Confirmability: Are the findings accurate? Can you find other ways to corroborate your results? These efforts are crucial in a rigorous qualitative study, not only to generate confidence in the results but also to reflect the truthfulness of the participants’ perspectives. Two of the most commonly applied strategies include audit trails and reflexivity.

Audit trails are likened to a research study blueprint, outlining detailed procedural records maintained by the primary researcher. This blueprint is accessible to an external researcher so s/he can attempt replication; if a study can be replicated with similar results, confirmability is strengthened. Designing an audit trail must include observations about how the procedures worked, and changes recommended for future study. An audit trail does not just tell another researcher what to do, but it also includes the researcher’s perspective on how the process worked.

IR Applications: In an IR office, audit trails are particularly valuable given the transience of staff and directors. There is no sense in reinventing the proverbial research wheel; a carefully outlined audit trail for recurring and large-scale studies is an especially important research practice.

By contrast, reflexivity is the incorporation of the researcher’s background, knowledge, bias, methodology, and perspective superimposed onto a study. Reflexivity represents what the researcher knows about himself and the participants, continuously recorded as a way to offset preconceived notions about the research which might interfere with data analysis and interpretation (Malterud, 2001). In other words, to what extent has the researcher worked to neutralize his or her own bias, motivation or interest as findings are reported?

IR Applications: As noted above, the goal of qualitative research is not to aim for generalizability but to seek depth, detail, and the multilayered perspective of each participant’s experience with the phenomena in question. For an institutional researcher, this reflexive process may be as simple as bracketing before interviews or focus group sessions, or it may be more complicated and involve the recording of field notes during an observation exercise.

Authenticity: This final element in the trustworthiness framework has been recently added to the discussion and is strongly endorsed by many qualitative researchers (Patton, 2002; Polit & Beck, 2011). This strategy focuses on the contextual purpose of the research, identifying the intended value of the research. How does the research benefit participants? Are all the realities represented to give meaning to the findings?

IR Applications: For an IR professional, authenticity must serve as the ultimate goal for any qualitative study in the campus setting. The intended value of a study, the implied benefit to many stakeholders, and the assurance that as many voices as possible were considered are all issues that must be addressed when presenting the findings to campus groups. Many of these issues are typically addressed in the executive summary, prefacing a report, or in the summary statements where these questions are ‘answered’. Additionally, customizing the results for different campus constituencies constitutes another way to increase the usefulness and applicability, and even the authenticity, of the findings.

Conclusion

This overview outlines the evaluative strategies to assess trustworthiness, which is arguably a complex concept. Many excellent resources are available and noted in the lists below; while Lincoln and Guba (1985) are most commonly cited, many other researchers have entered the conversation in recent years (Whittemore, Chase, & Mandle, 2001), offering important contributions and insights. Overall, researchers who do not believe there is any way a qualitative study can be as reliable, valid, believable, or useful as a quantitative study should be assured that when the principles of trustworthiness are diligently applied, a qualitative study is just as rigorous and valuable as any quantitative study. But as Morse, Barret, Mayan, Olson, and Spiers (2002) note: “While strategies of trustworthiness may be useful in attempting to evaluate rigor, they do not in themselves ensure rigor” (p. 9). The quest for rigor in any qualitative study ultimately resides with the quality of the researcher’s purpose and practice, and the verity of the unique depth and breadth of each participant’s ‘lived experience.’
REFERENCES


RESOURCES


Popay, J., Rogers, A., & Williams, G. (1998). Rationale and standards for the systematic review of qualitative literature in health services research. Qualitative Health Research. 8(3), 341-351.