

Trustworthiness

*Helen Foster-Collins, Department of Education
University of Exeter, UK
hf270@exeter.ac.uk*

Abstract

The terms validity and reliability have been contested within naturalistic research, with some attempts to redefine these concepts to increase their relevance for qualitative research. For my own research on informal workplace support for new teachers, I chose the alternative term ‘trustworthiness’ and additionally drew upon texts which discussed ‘rigour’ in qualitative contexts. I used the five strands outlined by Guba and Lincoln as a useful starting point for thinking about trustworthiness during my methodological design. In this paper, I outline these five criteria and how I have applied them in my own research. I also consider their strengths and limitations and briefly appraise the use of criteria-based measures of quality in general. I argue that reflection on trustworthiness of research is an ongoing process, comparable to that of the continued imperative to consider ethics in a responsive and unfolding way. Ensuring that we remain reflective during all stages of qualitative research may help to increase trust: in the overall quality of research, the results of data analysis, and any conclusions and recommendations based on these.

Introduction

It is important to conduct research studies of the highest quality, given that we draw conclusions and make recommendations for action on the basis of our findings. Whether that be for new interventions, changes to policy and practice, or informing theory, our research has real-world implications. Moreover, for academics and the public to have confidence in the results of research, they should be able to feel confident that it has been carried out rigorously. Practices such as double-blind peer review, where experts scrutinize the methodology, findings, and conclusions of research before publishing, can help build confidence. Such practices are difficult to carry out, however, if insufficiently detailed information is provided on how research was conducted. Finally, although qualitative research methods are used widely within social sciences, historically there have been questions over their value compared to quantitative methods (Lincoln and Denzin, 2000). For all these reasons, qualitative research communities should continue to engage in discussions about how best to promote quality.

Paradigmatic Assumptions Underlying ‘Quality’ Assessments

Most qualitative research is based within naturalistic, interpretivist paradigms (Guba and Lincoln, 1982), having the social world as its object, and using methods such as observation, interviews, and textual analysis to explore designated topics in all their complexity. Unlike quantitative research, it does not try to describe the world as it ‘really is’. There is no desire to manipulate variables, determine causal relationships, draw generalisations or make predictions (Golafshani, 2003). Instead, qualitative researchers see interpretation as inherent to our understandings of the world, with these interpretations situated within wider socio-cultural contexts of shared beliefs, meanings, and practices (Schwandt, Lincoln and Guba, 2007), and resting upon collaboratively constructed language concepts (Schwandt, 1994). It is therefore common for qualitative researchers to explicitly acknowledge possible influences on data production, such as social, cultural, and political climates, researcher-participant dynamics, and individual researcher standpoints. They also recognize that alternative interpretations of their data are possible.

A Criteria Approach to Ensuring Quality

However, it is still important to establish some guidance for quality and rigour if we are to avoid an ‘anything goes’ attitude towards qualitative research (Silverman, 2006). We have the same need to demonstrate that our work has value, that our findings and conclusions are ‘credible’, and that they genuinely answer the questions we set out to explore (Kline, 2008). One approach to addressing rigour has been to draw up sets of evaluative criteria specifically relating to qualitative research. The best-known of these criteria were set out by Guba and Lincoln (1981, 1989) to demonstrate what they called ‘trustworthiness’. Despite having been developed almost forty years ago, and notwithstanding the fact that Guba and Lincoln have since challenged their own criteria to some degree (Lincoln, 2010), this set of ‘standards’ is still widely used and cited. Each set of criteria was intended to correspond to accepted measures of quality in positivist paradigms and address similar concerns, but they were adapted for use within naturalistic paradigms. Four main strands of trustworthiness were initially described: credibility; transferability; dependability; and confirmability. A fifth element – authenticity – was

later added (Guba and Lincoln, 1986). These five aspects formed the main starting point for thinking about trustworthiness, and so I will both define these and illustrate how I have sought to apply them in my own research. I will also evaluate this approach to trustworthiness, and discuss how reflecting upon quality continues to challenge my understandings.

Informal Workplace Support for Early-Career Professionals

My own PhD research explores early-career professionals' experiences of workplace support during their first year, with a particular focus on informal support. These 'stories of support' are being analysed with the aim of discovering: who provides support; what types of support they give; and which socio-cultural and organizational factors influence this. Having previously analysed secondary data from first-year postgraduate doctors, my current focus is the collection and analysis of new narrative interviews from recently trained secondary-school teachers in their first year of practice.

Reflecting on the Five 'Criteria' of Trustworthiness

Table 1 below outlines each of the criteria of trustworthiness, as devised originally by Guba and Lincoln and further developed by others. It also includes illustrations of how I have implemented these and some reflections on that process.¹⁰

Table 1: Five criteria of trustworthiness as applied to narrative teacher interviews

	Criteria	Criteria	Reflection
Credibility	<p>Credibility is analogous to internal validity in quantitative research and seeks to answer the question: to what extent do the data and research findings show 'a true picture of the phenomenon under scrutiny'? (Shenton, 2004).</p> <p>Guba (1981) warned that data could be 'distorted' if researchers failed to challenge their own preconceptions and suggested that data could be influenced by factors such as a lack of rapport, or participants anticipating researchers' aims.</p> <p>Some suggested solutions have included: 'prolonged engagement' with participants and context, peer discussion and co-analysis, and rich descriptions of data collection contexts (Morrow, 2004).</p>	<p>In face-to-face interviews, I tried to develop rapport by bringing refreshments, using informal interview schedules, and describing how narrative methods aimed to capture teachers' own unique stories, with no right or wrong answers. I also reassured respondents regarding anonymity and confidentiality in the hope that this would enable them to speak more freely. In my reflective notes, I noted that rapport may be more difficult to achieve during telephone interviews due to the absence of body language cues.</p> <p>Active discussions with my supervisory team were used to try and uncover my own preconceptions about teacher support, such as how I might draw on my own experiences of working within education as a teaching assistant.</p> <p>Prolonged engagement is not usually deemed necessary in interview research (Morse, 2015).</p>	<p>The concept of 'rapport' has recently been problematized (Lincoln, 2010). It is a difficult thing to define or demonstrate to others but encompasses ideas such as empathy and respect. Within my study, it is perhaps better considered as a conscious intention to put participants at ease. Rapport may also be viewed as something which occurs not just during interviews but throughout all participant contacts (Miller, 2007). The idea of a 'true picture' of social phenomena 'independent of us that can be discovered' is suggestive of objectivity and realist ontologies (Smith and McGannon, 2018). However, I can reflect upon and acknowledge my own subjective perspectives where possible and aim to represent the varying perspectives of my teacher participants in research outputs.</p> <p>Given opportunities to ask questions, some participants did enquire about the wider aims of the research and what their data would be used for. However, more did not, giving me little access to their perceptions of the research as a whole.</p>

¹⁰ Due to limitations of space, it is not possible to include all of my reflections here.

	Criteria	Criteria	Reflection
Dependability	<p>Dependability is analogous to reliability in quantitative research, which is concerned with stability of findings (Anney, 2014). In qualitative research, it is not possible to replicate findings, but it may be possible to replicate methodology to an extent. Therefore, dependability in this case is chiefly concerned with research processes being transparent and 'auditable' (Guba and Lincoln, 1989).</p> <p>If research methods are explicitly described in sufficient detail, avoiding vague statements, e.g. that data 'was analysed' or 'themes emerged' (Anfara, Brown and Mangione, 2002), this allows others to critique the methods and conclusions drawn, and repeat methods if desired (Shenton, 2004; Aguinis and Solarino, 2019). This may include timelines of activities, notes on thematic development, and influences on data collection and analysis (Morrow, 2004). It is also desirable that published quotes be chosen to represent the variety, 'range and tone of responses' found in the data (Roberts, Priest and Traynor, 2006).</p>	<p>I am using Framework analysis, a type of thematic analysis described by Ritchie and Spencer (1994; 2003). This was primarily chosen for its transparent procedure, which is systematically applied to the whole data, and is easily referenced, understood, and repeated by others (Lacey and Luff, 2001).</p> <p>I have clearly described the research process undertaken, including the extent to which analysis was driven by data and previous theory. The number of quotes included in publications is necessarily constrained. However, they will be carefully selected to try and represent all major themes identified and the full variety of participant viewpoints.</p>	<p>Inevitably, there will be aspects of the research process which remain 'invisible' to others, given the messy complexity of data analysis. Additionally, difficult decisions may have to be made regarding what is shared, for instance, due to space limitations or considerations of confidentiality and anonymity. A high level of methodological detail will be shared in my thesis. This is more challenging when writing for journals, but supplemental data will be made available (e.g. my coding framework) via online appendices where possible.</p>

	Criteria	Criteria	Reflection
Confirmability	<p>Confirmability is analogous to objectivity in quantitative research and refers to whether one may be 'confident' in the products of research (Holloway and Wheeler, 2002) and whether findings plausibly represent the data and participant views (Graneheim and Lundman, 2004).</p> <p>Whilst acknowledging the emphasis on multiple possible perspectives in qualitative research, one would wish that findings were as representative as possible of 'the situation being researched' rather than the researcher's 'personal beliefs, motivations or pre-existing theories' (Gasson, 2004).</p> <p>Again, a thorough description of research processes is helpful, so readers may follow how results were obtained from the data (Bowen, 2009). Additionally, although complete neutrality is not possible, researchers can reflect upon their subjective viewpoints and assumptions to make biases more explicit (Koch, 2006). In addition, co-analysis and discussion with peers may provide 'congruence' (Elo et al., 2014), acting as a check against 'wild interpretations' of the data. 'Immersion' in the data and field may help researchers gain deeper understandings of context (Anney, 2005).</p>	<p>In my own research, a systematic process is being implemented, with an analytical framework being applied to the entirety of the data. This, it is suggested, may help to avoid 'cherry-picking'.</p> <p>Roberts, Priest and Traynor (2006) suggest that researchers move continually between the data, their interpretations, and back again, to check understandings. Therefore, the framework created will be left open to revision at all stages of data coding, so that novel themes, categories, and understandings may be identified. Additionally, active discussions will be carried out with my wider research team throughout all stages of the project, including sharing sections of raw data and corresponding analysis, so that my personal understandings and interpretations might be challenged.</p>	<p>There are some practical issues with co-analysis, with suggestions that power differentials can influence which interpretations of the data are privileged. It is not possible to claim that no power differentials exist within my supervisory relationships. However, I feel that in my research team, there is a good balance between experienced researchers acting as 'critical friends' (Costa and Kallick, 1993) and support to develop my own ideas.</p>
Transferability	<p>Transferability is comparable to the concepts of external validity or generalizability in quantitative research. Qualitative research is not 'generalizable' in the traditional sense, and so the results of research undertaken in one context are only 'transferable' to another context to the extent that they are similar. To make this assessment, the reader requires sufficient information regarding 'researcher as instrument' (Morrow, 2004), participants recruited, data collection contexts, and researcher-participant interactions (Shenton, 2004).</p>	<p>I have documented the number and broad location of schools involved, recruitment methods and inclusion criteria, participant demographics, the number, length, and timings of data collection episodes, interview questions and schedules, as well as the nature of workplace contexts studied.</p>	<p>There are tensions between the amount of information required to allow transferability judgements and space constraints in journals. Moreover, there may be limits to information sharing, due to ethical concerns over confidentiality and anonymity. In my own work, I have described the broad regions where data was collected and some participant demographics, whilst taking care to omit information which might inadvertently identify particular schools or participants.</p>

	Criteria	Criteria	Reflection
Authenticity	Guba and Lincoln's final aspect of trustworthiness is somewhat multifaceted but, broadly speaking, it addresses issues of fairness and ethical treatment of participants and stakeholders. This may encompass aspects such as: whether all viewpoints are equally represented; if participant feedback is acted upon; whether participants are informed as to how their data will be used; and whether the research empowers participants and stakeholders (Johnson and Rasulova, 2016).	A number of steps were taken to address such concerns in my own research. All teacher participants were informed as to the uses their data would be put in the information sheet . Events are also in planning for new teachers and educators to hear summaries of results and provide feedback on the practicality of proposed recommendations.	'Member-checking' in its truest sense – where raw transcripts are shared with participants for feedback – is not included in my research design. Although some researchers see this as an essential step, it can raise further methodological dilemmas (Carlson, 2010). For instance, how one resolves divergences of interpretation by participants and researchers, whether participants have time to actively engage with re-reading transcripts, whether participant perspectives on the data have changed since collection or are influenced by personal interests, whether power differentials encourage participants to defer to researchers' 'expertise' (Estroff, 1995), and whether clear evidence exists that such practices truly increase credibility (Thomas, 2017).

General Limitations of the Five Criteria

A general limitation of Guba and Lincoln's concepts is an apparent overlap between the five criteria. Various authors have also defined these criteria differently, making their own suggestions for the actions required to ensure quality. Despite these differences, the central tenets of trustworthiness appear broadly the same: providing thorough descriptions of methodological processes to allow transparency, and requiring that researchers engage in reflexivity at every stage. Some researchers reject the idea that unified guidelines can be formulated, due to the wide variations in qualitative research methodologies and lack of consensus on what these criteria should be (Rolfe, 2006; Sandelowski and Barroso, 2002). However, this is not a reason to abandon all thought of quality standards, given that academics frequently fail to reach consensus and yet continue to explore contested topics through lively debate. Instead, we might see quality criteria more as guiding concepts, which aid researchers' thought processes during the stages of design, analysis, and evaluation (Hammersley, 2005). A further approach to quality has been to develop 'lists' of characteristics deemed desirable in qualitative research, the idea being that these do not represent the last word but remain open to reinterpretation and reconstruction over time (Smith and Deemer, 2000). This approach has the advantage of allowing greater flexibility in how we implement quality, depending on the methodology and underlying epistemological stance, whilst providing constructive guidance for researchers and reviewers.

Further Reflections on Trustworthiness

The idea of a 'universal' set of criteria is highly appealing. This is particularly the case for new researchers, who may have sincere intentions to produce the highest standards of research but lack confidence in how to achieve this practically. This is exacerbated by a multitude of approaches which outline 'the way' to achieve quality research (Creswell and Miller, 2000). I would argue that using Guba and Lincoln's criteria was indeed a useful starting point for me as a researcher. However, critically reflecting on them and looking beyond to other conceptions of rigour continues to challenge my understandings of quality, and this emergent thinking can hopefully be integrated into my research practice. I would liken this process of thinking about trustworthiness to how we consider ethics in research. The use of ethics forms, ratified by academic institutions, can encourage a way of thinking in new researchers that ethics is a once-only process. A hurdle which one must 'get over' and then never return to, rather than a set of principles and continued self-reflection which applies to the entirety of the research process (Guillemin and Gillam, 2004).

The Ethics of Trustworthiness

In this regard, I was inspired by a presentation by Brannelly and Barnes (2018), who put care at the centre of their research ethics. Drawing on Tronto (1993; 2013), they argued that caring for participants requires a degree of attentiveness, responsibility, competence, responsiveness, and solidarity with participants. The idea of caring for one's participants is somewhat abstract, yet we have all read papers in which the researcher's passion for their topic is clear and conveys a desire to represent their participants. When thinking about trustworthiness, this is not simply a process of laying out a trail of the research path we have taken for others to examine and critique, though clearly this is important. It is an ethical process which involves asking ourselves some piercing questions: Why am I doing this research? Who is it for? Am I telling the stories that participants want to be told? Relating this to the ethics of care, I can see that in order for my research to be trustworthy, this involves aspects of caring for participants and my research, ensuring my research skills are sufficient to do the data justice, recognising the varying aims and needs of participants, and remaining responsive to feedback. That is not to say that I can meet every need of every participant in every way; their views will invariably come into conflict with those of other participants and my own, acting as researcher-interpreter for their stories. In my own research, I face difficult choices regarding which of the many findings are most relevant to publish and which participant quotes should be selected to illustrate these. I am therefore required to ask myself: Am I fully conscious of any motives and interests which might influence this process? Have I favoured findings which support my own beliefs, are more interesting to me personally, come from a 'favoured' participant, incite controversy or play it safe, or for which it is easier to suggest solutions? That is not to say that we should present results which are vague or boring, that court or avoid controversy, or that difficult problems be presented with no apparent solutions (except that 'more research is needed'). However, participants entrust us with their personal experiences and stories for a reason. Therefore, it is important that we approach the research process with a clear intent, to ensure our participants' voices are heard and that we try to represent their best interests.

Final Conclusions

As a relatively new researcher, my thinking around trustworthiness is continually developing. Guba and Lincoln's criteria offered a useful guide during my research design, but perhaps should not be viewed as being 'set in stone'. Alternatively, 'list' approaches to rigour may allow greater flexibility when identifying which characteristics of trustworthiness apply to particular methodologies, topics, and settings. Furthermore, the use of criteria does not absolve us from deeper reflections on whether we can go above and beyond them; to imbue our research with genuine concern for participants and the impact of our research. In short, that as researchers, we are indeed 'trustworthy'.

References

- Aguinis, H. and Solarino, A. M. (2019) Transparency and replicability in qualitative research: The case of interviews with elite informants. *Strategic Management Journal*.
- Anfara Jr, V. A., Brown, K. M. and Mangione, T. L. (2002) Qualitative analysis on stage: Making the research process more public, *Educational Researcher*, 31(7), 28–38.
- Anney, V. N. (2014) Ensuring the quality of the findings of qualitative research: Looking at trustworthiness criteria, *Journal of Emerging Trends in Educational Research and Policy Studies (JETERAPS)*, 5(2), 272–281.
- Bowen, G. A. (2009) Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40.
- Brannelly, T. and Barnes, M. (2018) Brannelly Barnes Ethics of Care and Research, in *NCRM Research Methods Festival 2018, 3–5 July 2018*, University of Bath.
- Campbell, J. L., Quincy, C., Osseman, J. and Pedersen, O. K. (2013) Coding in-depth semi-structured interviews: Problems of unitization and intercoder reliability and agreement, *Sociological Methods and Research*, 42(3), 294–320.
- Carlson, J. A. (2010) Avoiding traps in member checking. *The Qualitative Report*, 15(5), 1102–1113.
- Costa, A. L. and Kallick, B. (1993) Through the lens of a critical friend. *Educational Leadership*, 51, 49–51.
- Creswell, J. W. and Miller, D. L. (2000) Determining validity in qualitative inquiry, *Theory into Practice*, 39(3), 124–130.
- Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K. and Kyngäs, H. (2014). Qualitative content analysis: A focus on trustworthiness, *SAGE open*, 4(1), doi:10.1177/2158244014522633.
- Estroff, S. E. (1995) Whose story is it anyway? Authority, voice, and responsibility in narratives of chronic illness, in S. K. Toombs, D. Barnard and R. A. Carson, eds, *Chronic Illness: From Experience to Policy*, Indiana University Press, Bloomington, IN, pp. 77–102.
- Gasson, S. (2004) Rigor in grounded theory research: An interpretive perspective on generating theory from qualitative field studies, in M. E. Whitman and A. B. Woszczyński, eds, *The Handbook of Information Systems Research*, IGI Global, Hershey, PA, pp. 79–102.
- Golafshani, N. (2003) Understanding reliability and validity in qualitative research, *The Qualitative Report*, 8(4), 597–606.
- Graneheim, U. H. and Lundman, B. (2004) Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness, *Nurse Education Today*, 24(2), 105–112.
- Guba, E. G. and Lincoln, Y. S. (1981) *Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and Naturalistic Approaches*, Jossey-Bass, San Francisco.
- Guba, E. G. and Lincoln, Y. S. (1982) Epistemological and methodological bases of naturalistic inquiry, *ECTJ*, 30(4), 233–252.
- Guba, E. G. and Lincoln, Y. S. (1989) *Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and Naturalistic Approaches*, Jossey-Bass, San Francisco.
- Guillemin, M. and Gillam, L. (2004) Ethics, reflexivity, and ‘ethically important moments’ in research, *Qualitative Inquiry*, 10(2), 261–280.
-

-
- Hammersley, M. (2005) Assessing quality in qualitative research, *ESRC TLRP Seminar Series: Quality in Educational Research*, July, 103–113.
- Holloway, I. and Wheeler, S. (2002). Ensuring trustworthiness and quality, in I. Holloway and S. Wheeler, eds, *Qualitative Research in Nursing*, 2nd edition, Blackwell, Oxford, pp. 250–263.
- Johnson, S. and Rasuloova, S. (2016) *Qualitative Impact Evaluation: Incorporating Authenticity into the Assessment of Rigour* (No. 45), Bath Papers in International Development and Wellbeing.
- Kline, W. B. (2008) Developing and submitting credible qualitative manuscripts, *Counselor Education and Supervision*, 47(4), 210–217.
- Koch, T. (2006) Establishing rigour in qualitative research: The decision trail, *Journal of Advanced Nursing*, 53(1), 91–100.
- Lacey, A. and Luff, D. (2001) *Qualitative Data Analysis*, Trent Focus Group, Nottingham.
- Lincoln, Y. S. and Guba, E. G. (1986) But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation, *New Directions for Program Evaluation*, 1986(30), 73–84.
- Lincoln, Y. S. and Denzin, N. K., eds. (2000) *The Handbook of Qualitative Research*, Sage, London.
- Lincoln, Y. S. (2010) ‘What a long, strange trip it’s been...’: Twenty-five years of qualitative and new paradigm research, *Qualitative Inquiry*, 16(1), 3–9.
- Miller, T. (2017) Telling the difficult things: Creating spaces for disclosure, rapport and ‘collusion’ in qualitative interviews, *Women’s Studies International Forum*, 61(March), 81–86.
- Morrow, S. L. (2005) Quality and trustworthiness in qualitative research in counselling psychology. *Journal of counselling psychology*, 52(2), 250.
- Morse, J. M. (2015) Critical analysis of strategies for determining rigor in qualitative inquiry, *Qualitative Health Research*, 25(9), 1212–1222.
- Ritchie, J. and Spencer, L. (1994) Qualitative data analysis for applied policy research, in A. Bryman and R. G. Burgess, *Analyzing Qualitative Data*, Routledge, London, pp. 173–194.
- Ritchie, J., Spencer, L. and O’Connor, W. (2003) Carrying out qualitative analysis, in J. Lewis and J. Ritchie, *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, Sage, London, pp. 219–262.
- Roberts, P., Priest, H. and Traynor, M. (2006) Reliability and validity in research, *Nursing Standard*, 20(44), 41–45.
- Rolfe, G. (2006) Validity, trustworthiness and rigour: Quality and the idea of qualitative research, *Journal of Advanced Nursing*, 53(3), 304–310.
- Sandelowski, M. and Barroso, J. (2002) Reading qualitative studies, *International Journal of Qualitative Methods*, 1(1), 74–108.
- Schwandt, T. A. (1994) Constructivist, interpretivist approaches to human inquiry, *Handbook of Qualitative Research*, 1, 118–137.
- Schwandt, T. A., Lincoln, Y. S. and Guba, E. G. (2007) Judging interpretations: But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation, *New Directions for Evaluation*, 2007(114), 11–25.
- Shenton, A. K. (2004) Strategies for ensuring trustworthiness in qualitative research projects, *Education for Information*, 22(2), 63–75.
-

-
- Silverman, D. (2006) *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction*, Sage, London.
- Smith, J. K. and Deemer, D. K. (2000) The problem of criteria in the age of relativism, *Handbook of Qualitative Research*, 2, 877–896.
- Smith, B. and McGannon, K. R. (2018) Developing rigor in qualitative research: Problems and opportunities within sport and exercise psychology. *International review of sport and exercise psychology*, 11(1), 101-121.
- Thomas, D. R. (2017) Feedback from research participants: Are member checks useful in qualitative research?, *Qualitative Research in Psychology*, 14(1), 23–41.
- Tronto, J. C. (1993) *Moral Boundaries: A Political Argument for an Ethic of Care*, New York University Press, New York.
- Tronto, J. C. (2013) *Caring Democracy: Markets, Equality, and Justice*, New York University Press, New York.