Criteria for evaluating qualitative studies

Developed by the Qualitative Research and Health Working Group, LSTM (Helen Bromley, Grindl Dockery, Carrie Fenton, Bertha Nhlema, Helen Smith, Rachel Tolhurst, Sally Theobald) in 2002

Introduction

This checklist has been designed for people with a basic understanding of qualitative research methodology. Increasing recognition of the importance and benefits of using qualitative methods in health research has triggered a demand for a means of assessing the quality of qualitative studies. For each criteria there is an assessment scale, which can be used to guide evaluations of many studies; alternatively the criteria can be used alone or by writing your own comments in the comments box.

Specific terminology has been developed in the field of qualitative research, some of which is arguably parallel to terminology used in quantitative research. However, there is some disagreement between qualitative researchers about the usage and meaning of different terms.

Some assessment criteria should ideally be applied to the entire research cycle or several different phases, but have been included only at a particular point(s) to include repetition. For example, the concepts of 'trustworthiness' refers to an assessment of the entire research process, but has been included in the checklist at the findings/interpretation stage because this is generally where an assessment of trustworthiness is often made in a research report or paper. The checklist should be used flexibly to suit the purposes of the user – for example an assessment of a research ethics proposal would involved consideration of assuring trustworthiness at the planning stage.

This is a preliminary checklist and will be revised; please feedback to the group your comments and experience of using the checklist.

1. Theoretical/ epistemological issues	Assessment (circle most relevant)	Comments
(a) Is a qualitative approach appropriate to answer the research question?	Appropriate Unclear Inappropriate	
(b) Is a qualitative approach justified by the author? >>HINTS: Does the research methodology seek to understand or illuminate the subjective experiences or views of those being researched? Does the research methodology seek to understand WHAT is happening and the reasons WHY observed situations, outcomes or discourses occur?	Justified Unclear Not justified	
(c) Is the purpose of conducting the research adequately described and justified	Justified Unclear	

>>HINTS	Not justified	
Was a review of the secondary data conducted and is it presented?		
Is the research linked to policy or practice development processes?		

2. Study design	Assessment (circle most relevant)	Comments
(a) Is the context of the research adequately described?	Adequate Unclear Inadequate	
(b) Is the research question relevant to the context described?	Relevant Unclear Not relevant	
(c) Are the research aims/objectives/questions clearly defined and focused?	Adequately defined Unclear Poorly defined	
(d) Are the methods used appropriate to the research question? >>HINTS: Is a range of methods used for triangulation, or is use of a single method justified? Do the methods investigate what they claim to? Have the best methods been chosen to address the research question?	Appropriate Unclear Inappropriate	

3. Sampling and data collection	Assessment (circle most relevant)	Comments
(a) Is the sampling strategy appropriate to the research question?	Appropriate Unclear	

>>HINTS: Usually purposive or theoretical NOT random or representative. Is the sample sufficient to understand the study context and population? Was the sampling pre-determined or did it evolve as the fieldwork progressed?	Inappropriate	
(b) Is the choice of sampling strategy justified? >>HINTS: Are the reasons for this choice discussed/compared to other strategies? Who was selected and why? (consider gender, age, ethnicity, marital status) How were participants selected and why? Is it clear why some participants were not selected?	Adequately justified Unclear Not justified	
(c) Are data collection procedures clearly described? >>HINTS: How was data collected? (topic guides, checklists) Were data collection tools pilot tested? Where was data collected and why was this location chosen? (privacy, confidentiality, familiarity) How was the data recorded and why? (tape recorded, notes)	Clear Unclear	
(d) Are the roles of researchers clearly described? >>HINTS: Who conducted the research, how were they selected? Are the researchers skills, motives, background, position in terms of power-relations (gender, age, ethnicity, employment relations etc.) and perspective described and discussed?	Clear Unclear	
(e) Are ethical issues addressed in data collection and adequately discussed? >>HINTS: How was the research explained to the participants?	Adequate Unclear Inadequate	

What consent procedures were used?		
How were confidentiality and privacy assured?		

4. Analysis	Assessment (circle most relevant)	Comments
(a) Is the data analysis procedure explicit? >>HINTS: Is it clear how the researcher processed the raw data to arrive at the stated results? Were the categories and themes identified in advance, or derived from the data? Are all data taken into account in the analysis? Are responses/experiences compared and contrasted across different groups/individuals/study sites?	Explicit Unclear Vague	
(b) Is the data analysis procedure reliable/dependable? >>HINTS: Who was involved in the analysis and at what stage? Did more than one person (including researchers and other stakeholders) identify themes and code transcripts?	Reliable Unclear Potential bias	

5. Findings/interpretation/trustworthiness	Assessment (circle most relevant)	Comments
(a) Are the findings valid/internally coherent/trustworthy >>HINTS: Are findings drawn from analysis of collected data rather than the researcher's preconceptions? Is there adequate critical discussion for and against the researcher's arguments? E.g. Are negative and divergent views adequately discussed? Are quotes used to substantiate the researchers conclusions from the analysis? Is triangulation or data cross-checking used? Have findings been validated by respondents?	Valid Unclear Invalid/ potential bias	

Has the researcher critically reflected on his/her own bias, role and influence?		
Has the research critically reflected on the quality of the data collected and skills of the research team?		
(b) Are the findings relevant? >>HINTS: Are the findings relevant to the study aim/objectives/questions? Do they contribute new knowledge or understanding? How important are the findings in local context? (geographical, cultural, political, socio-economic)	Relevant Unclear Limited relevance	

6. Implications/limitations	Assessment (circle most relevant)	Comments
(a) Are the implications of the study clearly defined? >>HINTS: Are the findings placed in local context? (geographical, cultural, political, socio-economic) Have findings been disseminated to key stakeholders including participants? Are the findings discussed in wider context? (in relation to other studies on the same topic) Are recommendations made for policy and practice?	Clear Unclear	
(b) Is there adequate discussion of the study limitations? >>HINTS: Are study limitations described and accounted for? (cost, time, resources) Are the weaknesses of the study design discussed?	Adequate Unclear Inadequate	

Acknowledgements

BSA Medical Sociological Group (1996) Criteria for the evaluation of qualitative research papers. Medical Sociology News. 22.

Cobb, A.K, & Hagemaster, J.N. (1987) Ten criteria for evaluating qualitative research proposals. Journal of Nursing Education. 26:4; 138-143.

Cochrane Qualitative Methods Network. 10 questions to help you make sense of qualitative data. Critical Appraisal Skills Programme (CASP) 1998. (Accessed via website: ???

Crombie, I.K. (2000) The pocket guide to critical appraisal. BMJ Publishing Group. London. UK.

Greenhalgh, T. & Taylor, R. (1997) How to read a paper: papers that go beyond numbers (qualitative research). BMJ; 315:740-743

Malterud K. Qualitative research: standards, challenges and guidelines. Lancet 2001;358:483-88.

Mays, N. & Pope, C. (2000) Qualitative Research in health care: Assessing quality in qualitative research. BMJ 320; 50-52.

Group contacts:

Helen Bromley, Lecturer LSTM, <hbromley@liv.ac.uk>

Grindl Dockery, Honorary Resaerch Fellow, LSTM, <grind_doc55@yahoo.co.nz>

Carrie Fenton, Research Associate, LSTM, <cmfenton@liv.ac.uk>

Bertha Nhlema, Researcher Equi TB Knowledge Prgramme, LSTM/Malawi, <bkisu2001@yahoo.co.uk>

Helen Smith, Research Associate, LSTM, <cjdhel@liv.ac.uk>

Sally Theobald, Lecturer LSTM, <sjt@liv.ac.uk>

Rachel Tolhurst, Research Associate, LSTM, <r.j.tolhurst@liv.ac.uk>

Evaluating Qualitative Research*

I. Criteria

A. Is substantive & formal theory produced and developed? Is the research question clear?

B. Has the work contributed usefully to knowledge?

C. Are the claims consistent with the data collected?

D. Are the results of the analysis credible to the participants and to readers?

E. Were the researchers competent to study people of the participants' culture? (Culture broadly defined.)

F. Can the findings be transferred to other settings? How clearly as the context or setting described (to permit this evaluation)?

G. How heavily were the findings influenced by the research method?

II. Increasing Validity

A. Triangulation

1. Comparing results from 2 or more methods

2. Comparing results from 2 or more informants

B. Respondent Validation

Checking the investigators' account with the subjects themselves, and then incorporate reactions into data.

C. Clearly describe methods of data analysis & collection

D. Make biases clear and examine how these may have affected data.

E. Deviant case analysis

Find contradictory data and refine analysis until can explain all or almost all cases.

F. Incorporate wide range of different perspectives; sample the full range of cases

* This discussion draws heavily on Mays & Pope (2000). Assessing quality in qualitative research. *British Medical Journal, 320,* 50-52.

Evaluating Information and Research

Criteria for evaluating qualitative research

- Beyond truth and objectivity
- Credibility
- Transferability
- Dependability
- Confirmability

Beyond truth and objectivity

Research has traditionally been associated with the a number of key criteria:

Truth - the idea that the object of research offers an accurate description or model of the the object of the study.

Objectivity (or neutrality) - the idea that the person of the researcher, with her biases, interests and particular experiences should not affect the research findings.

Applicability (or generalisability) - the idea that the findings of the research should apply in other contexts, either in the ability to predict, or in the sense that the findings from the particular research sample can be generalised to a broader population.

In addition one could also ask how well the particular research project has been designed and carried out; this is usually referred to as reliability or consistency and associated with the idea that the research is repeatable or replicable. These have provided the criteria by which much of scientific and social scientific research has been judged.

Although these might seem to embody a useful set of values, many researchers working from a more qualitative or interpretive perspective have argued that they not appropriate to all forms of inquiry. The idea of truth, which implies a relatively fixed and predictable external world, is suited to scientific research, but in contexts where one is dealing with people and with the social and cultural world it becomes a more problematic concept. The idea of objectivity has come under close scrutiny from a number of perspectives. Some feminist researchers for example have pointed to the considerable importance of the person doing the research, his or her impact on the dynamics of an interview, and indeed the questionable desirability of neutrality in contexts where research might be seen as a lever for social change. Similarly, whilst applicability might be a desirable goal in many situations, qualitative researchers often emphasise what is unique rather than what is representative.

However, this is not to say that one should operate without any guiding criteria, but that researchers need to generate new ways of assessing the values and standards by which their work should be judged. The aim of this section is to look at some of the criteria that might be considered appropriate to qualitative research in art, design and media. In order to do this I draw on the categorisation used by Lincoln and Guba 1985, and in particular the discussion of this

and qualitative criteria more generally in Seale 1999 [1]. Lincoln and Guba identify four criteria: credibility; transferability; dependability; and confirmability. With some adaption to meet the specifics of the particular research being discussed, I believe these are a useful set of criteria with which to think about research in art, design and media.

Credibility

The notion of credibilty recognises that when we are dealing with research in human settings there is no single correct answer, instead there are always multiple perspectives. The question is not whether the research offers the correct version of events, but whether the account is credible. The corollary of this is that there may be more than one credible account depending on the perspective one adopts. In fields such as art and design history for example this is non-controversial, the issue is not whether one offers the right interpretation of an image but whether you offer a credible interpretation for which there is supporting evidence. The researcher may be called to arbitrate between different possible arguments and explain why they choose one account in preference over another.

Lincoln and Guba argue that credibility builds up over time. If research over an extended period by many researchers points to the same conclusion or interpretation then it can be held to be more credible, though not unassailable. Put simply what the notion of credibility asks is how are the findings of this particular project consistent or inconsistent with other findings or the perspectives of alternative actors. If they are inconsistent, then how might this inconsistency be explained? What this means of course is that if you are putting forward conclusions that radically challenge existing ideas, you will be required to provide stronger evidence than if you are arguing in support of them.

Credibility can be demonstrated in other ways too, for example by testing your interpretation against that of others. Importantly, in qualitative research it is often valuable to test your conclusions with those you are researching. If we take the case of the interpretation of imagery, one could seek the views of those who produced the images, or perhaps better still, those who view the images, as a means of testing your own interpretation. Similarly, what are referred to as 'member checks' or 'respondent validation' can be used when you are doing research based on interviews or observation - do people recognise your view of their reality?

The notion of credibility implies a greater tolerance of ambiguity; the accounts produced by researchers cannot be considered as the final word on the matter, but are always open to question and modification on the basis of later evidence and argument.

Transferability

One test of research findings is the extent to which they can be applied to other contexts, or whether they have predictive power, this is sometimes referred to as external validity. However, qualitative research often deals with unique settings and the special case. By its very nature this kind of purposive sampling does not lend itself to generating findings that can easily be applied to other cases.

The criteria of transferability switches around the responsibility. Whilst the notion of generalisability suggests that researchers should work to ensure that the results they get can be extended beyond the sample they actually considered - this is typically the case with quantitative research based on statistical samples - the notion of transferability argues instead that researchers should offer a rich and detailed account of the specific circumstances of their research in order that others may be able to pick up some of the insights that are offered and use them in different contexts.

This of course should not absolve the researcher from thinking about the wider implications of their findings and the limits to their applicability, but modifies this expectation from being an absolute test of the research, to something more realistic in terms of qualitative work. Do you provide sufficient detail for other researchers to understand the specific circumstances of your research? What might others learn from your research and in what contexts might your findings offer some insight?

Dependability

In qualitative contexts it often makes little sense to think in terms of a research project being replicated by others. In this case Lincoln and Guba propose the notion of dependability. What this means in effect is the researcher making the process of the research itself open and accountable, a process they refer to as 'auditing'. This involves keeping a record of the major actions and decisions of the research process, how and why you chose your particular focus, what methods you used to collect data, and so on. A good research report will provide sufficient evidence for the reader to assess its dependability.

Of course accountability, does not guarantee dependability, it merely opens it up to a process of questioning. Have you considered a sufficient range of arguments in coming to your conclusions? Does the amount of first-hand investigation you have done warrant the statements you make? Is your choice of case studies appropriate to your aims?

Confirmability

The final criteria proposed by Lincoln and Guba is confirmability. This is offered in the place of the criteria of objectivity or neutrality that is a feature of conventional research evaluation. As I have suggested above, for many researchers outside of the sciences complete neutrality is neither possible, nor necessarily desirable. In many forms of social and cultural research the very choice of subject matter implies some kind of value judgement. Rather than obscure these values, as some have argued is what in effect happens in many positivist approaches to research, it is suggested these too should be reflected in the reporting of research.

Researchers should be self-reflexive about their impact on the research process, how who they are might facilitate their access to certain settings, how it might have impacted on the kind of responses they got whilst interviewing, and so on. Where the lack of neutrality might be seen to detract from the research findings, then what strategies were used to cope with this? As with the criteria of dependability, this involves some form of accountability or auditing.

What all four criteria suggest is that good qualitative research is about making an argument as to how and why you chose to do what you did, subjecting your own findings to external testing and critical reflection, and offering a transparent account of how it was done.

1. Lincoln, Y.S. & Guba, E.G. (1985) *Naturalistic Inquiry*. London: Sage; Seale, C. (1999) *The Quality of Qualitative Research*. London: Sage. Whilst this is in my view a very useful and clear introduction to the issues, there are a variety of different positions and many points of dispute.

Previous   Support   Home   Contents   Next