

Ensuring Rigor in Qualitative Research Within the Field of Cross-Cultural Psychology

Carolyn Demuth (cdemuth@uni-osnabrueck.de)

University of Osnabrueck, Faculty of Human Sciences, Institute for Psychology
Culture & Development, Artilleriestrasse 34, D- 49076 Osnabrueck, Germany

Abstract

Within recent years, there has been an increasing call for qualitative research in cross-cultural psychology. Despite this general openness, there seems to be some confusion about how to evaluate the quality of such research. This has been partly due to the heterogeneity of the field and the epistemological underpinnings of qualitative research that do not allow for standard criteria of rigor as in the traditional psychological research. Nevertheless, there is an emerging canon of recognized standards of good practice in qualitative research which the present paper will briefly discuss. The paper aims at motivating cross-cultural psychologists to produce high quality qualitative research that will contribute to the further advancement of the field.

Cross-cultural psychology has long been associated with a nomothetic approach that attempts to identify general laws and causal explanations following the model of natural science and using standard quantitative research methods (Boesch & Straub, 2006; Kim, Park, & Park, 2000; Ratner & Hui, 2003). This picture has been challenged in recent years as there has been an increasing call for qualitative research in cross-cultural psychology that goes hand in hand with a theoretical rapprochement towards various branches of cultural psychology. This trend has also been visible for instance in the title of the XIX International Congress of IACCP in 2008 in Bremen (*“Crossing Borders – (Cross-)cultural psychology as an interdisciplinary multi-method endeavor”*), in an increasing number of qualitative research papers presented over the last IACCP congresses, the two ARTS workshops on qualitative (and mixed) methods, and finally in the publication of a JCCP special issue on qualitative and mixed methods in 2009 (Karasz & Singelis, 2009).

Despite the effort of a number of scholars to integrate qualitative methods into the field of cross-cultural psychology, there seem to exist some misconceptions about how to conduct qualitative research properly in order to meet empirical standards of rigor. Actually, very few studies in the field go beyond explorative procedures and use established methods that have been developed within the qualitative paradigm. Some use ‘open procedures’ such as open-ended questions in questionnaires, (semi-)structured interviews, or video recordings but in principle remain within the quantitative paradigm as far as the logic of analysis (and of the assessment) of the data are concerned. While this misconception of qualitative research is certainly partly due to the lack of training at universities where qualitative methods are still not taught as part of the regular curriculum in psychology, it is still surprising since there are sufficient handbooks available on the market and an increasing number of trainings and workshops are offered throughout the US and many parts of Europe.

Different Methodologies Require Different Criteria for Evaluation

Have qualitative methods finally arrived in cross-cultural psychology? I think not yet fully. Without going into a debate on the different epistemological underpinnings of quantitative and qualitative methods, I think it is essential to understand that qualitative procedures (I prefer to speak of ‘procedures’ rather than of ‘methods’ since qualitative research generally does not draw on standardized techniques) follow a different logic than quantitative methods. We are not simply talking about different methods in the sense of different techniques but of different *methodologies*. This is important to understand if we want to evaluate the quality of empirical research. It is surprising that still sometimes erroneously traditional criteria of rigor developed within the quantitative paradigmatic framework are applied to evaluate qualitative research (e.g., representative sample, sample size) without realizing the inappropriateness of these criteria for these procedures. This is even more surprising as specific criteria for the evaluation of qualitative research have been developed and published within the past

two decades (e.g., Drisko, 1997; Elliot, Fischer, & Rennie, 1999; Flick, 2009; Gaskell & Bauer, 2000; Guba & Lincoln, 1982; Lamnek, 2006; Lincoln & Guba, 1985; Patton, 2002; Richardson, 1996; Seale & Silverman, 1997; Silverman, 2001, 2005; Steinke, 2004; Stiles, 1993).

While some criteria of good practice of scientific research apply to both quantitative and qualitative research, others apply specifically to qualitative approaches (Elliot et al., 1999; Silverman, 2001). General criteria are: relationship of the study to relevant literature, clarity of research questions, methodological appropriateness (choice of method, choice of sampling strategy), informed consent and ethical research conduct, specification of methods, appropriately tentative discussion of implications of research data and understandings, clarity of writing and contribution to knowledge (Elliot et al., 1999).

Qualitative research is of course in itself a very heterogeneous field that embraces a great variety of approaches that have been developed within very diverse theoretical and philosophical frameworks. This diversity, together with the fact that qualitative research does not follow standardized procedures, makes it difficult to define a common set of criteria that ensure rigor and quality of good practice. In the past, different research traditions have developed a variety of rules based on the logic of the relevant approach. Specific criteria for particular qualitative procedures have for instance been developed for the evaluation of Grounded Theory studies (Corbin & Strauss, 1990; Strauss & Corbin, 1990) and for the evaluation of discourse analysis studies (Potter, 1996, 2007; Coyle, 2000; Taylor, 2001a; Phillips & Jorgensen, 2001).

While there is no consensus about the best criteria for evaluating qualitative research, there has been an emerging canon of a potential set of broadly defined criteria that generally apply to qualitative research (and that need to be further specified for the concrete research project; Elliot, Fischer, & Rennie, 1999). These criteria are meanwhile commonly discussed in text books on qualitative research (Camic, Rhodes, Yardley, Rhodes, & Yardley, 2003; Y. S. Lincoln, 1995; Y. S. Lincoln & Guba, 1985; Lyons & Coyle, 2007; Mey & Mruck, 2010; Smith, 2008; Steinke, 2004; Willig, 2008).

Principles of Qualitative Research

In what follows, I would like to give an overview of the canon of good practice in qualitative research. In so doing, I hope to contribute to a state-of-the-art quality in qualitative cross-cultural research. In order to understand the logic that these criteria are based on, I consider it very helpful to keep in mind the basic principles that qualitative research – despite its diversity - is based on (e.g., Mey, 2010):

The Principle of Openness

The research design needs to be arranged in a way that allows the participants to lay open their personal subjective views and to behave as in every-day life. This principle applies to the formulation of a research question, to the sampling procedure, as well as to the actual analysis: no ex-ante hypotheses or pre-defined answers (as in a questionnaire) are to be formulated because the researcher then runs the risk of staying within his or her own perspective and expected categories. Theoretical knowledge is used as ‘theoretical sensitivity’ (Glaser, 1978) which helps the researcher to be aware of the subtleties of meaning of data. Sampling should be based on theoretical considerations (“theoretical sampling”; Glaser & Strauss, 1967; Strauss, 1987) rather than on the representativeness of a sample. This implies that the beginning of research is based on an initial sampling that comprises a small number of cases and allows to get first insights into the phenomenon under study. As the analysis proceeds, further cases are systematically added to include variation into the sample and to further develop the rudimentary theory/theoretical sketch. This is done by selecting new cases that allow for a minimum and maximum comparison with previous cases. Sampling thus follows an iterative rather than a linear process and continuously switches between data assessment and data analysis. Similarly, analysis follows inductive procedures deriving concepts and categories from the data in a recursive process rather than having a pre-defined set of categories that are applied to the material.

The Principle of Foreignness

This principle claims that the researcher is to refrain from pre-mature interpretations of the participant's utterances/behavior. The world-view of the participant is to be considered 'foreign' to the researcher and as something that still needs to be explored and discovered (this applies of course particularly to cross-cultural research, but also to any other research). The aim is to re-construct the subjective view or the meaning of the part of the participant in the sense of *Verstehen* as outlined by Weber (1968).

The Principle of Communication

This principle acknowledges that all data assessment involves a process of communicative interaction between the researcher and the participant which contributes to the understanding of the situation. This implies that the researcher is inevitably part of the co-construction of the data. The researcher therefore needs to critically reflect his role in the research process and interpret the participant's utterances and behavior as co-constructed.

Recognized Standards of Good Practice in Qualitative Research

Criteria for evaluating qualitative research need to be understood against the backdrop of these basic principles. The following criteria of evaluation can be considered as recognized standards of good practice in qualitative research:

Credibility

Credibility refers to the degree to which the findings reflect the actual *Lebenswelt* ("lived experience") of the participants. High credibility is therefore achieved by designing the study as non-intrusive as possible and by building a trustworthy atmosphere to assure that participants act in a way that they usually would.

Transparency (Intersubjective Traceability)

Transparency refers to the degree to which the way the researcher comes to his or her conclusions is made transparent to others and hence open for evaluation. Transparency is achieved by providing a clear documentation of the sampling strategy, data collection, transcription conventions, the individual steps of analysis, as well as the documentation of changes made to the research design. Moreover, auditability (i.e., ability to be audited) must be ensured through the availability of raw data (e.g., transcripts, video recordings, field notes), as well as the clear documentation of the process of analysis (individual steps of analysis as well as memos on preliminary ideas and hypotheses) and of the transcription notations that were used. For discourse analytical studies, sufficiently detailed transcription of talk in interaction according to conventional transcription rules in conversation analysis and discourse analysis (Heritage, 1984; Jefferson, 1984) must be provided that allows detection of how speakers orient to each other, and to reconstruct latent structures of meaning (Flick, 2009; Seale & Silverman, 1997, for demonstration of enhanced reliability and validity through detailed transcripts). Transparency also includes that the investigator discloses his or her expectations and preconceptions, and assumptions for the study (based on the researchers' knowledge and personal experience), before starting the analysis.

High transparency is, for instance, achieved when a detailed account of how the data were initially coded is provided, along with a report on how codes were modified through comparison of all instances and discussions between the researchers. An example of low transparency would be if there is little description is provided of how themes were identified and no checks on their consistency are reported (Smith, 2009).

Grounding of the Interpretation

This criterion refers to the degree to which interpretations are sufficiently grounded in the data. This can be achieved by

- (1) intensive engagement with the material and iterative cycling between observation (i.e. reading and re-reading transcripts, replaying audio and video material) and interpretation
- (2) providing sufficient text sequences as well as "thick descriptions" (Geerts, 1973) to support the interpretation. For codifying procedures, for instance, it must be demonstrated, how categories and concepts are generated and grounded in the data. Moreover, for Grounded Theory studies, it must be demonstrated that

concepts are systematically related and that a dense network of conceptual linkages has been developed. For conversation analytical and discourse analytical studies, this means that transcripts are presented along side its analytic interpretation to allow the reader to act as an auditor and to provide evidence for and against the researcher's arguments (Potter, 1996, 2004, 2007). Justification of claims is not primarily achieved by frequency of occurrence but by providing evidence for a *lawful pattern* in the way a certain structure is achieved (Bohnsack, 2001). Relevance of a phenomenon is therefore established by building aggregates of single instances and finding typical patterns of interaction, showing the *how* and the *what* rather than the *how often*. Some authors suggest, however, that frequencies of occurrences of these lawful patterns can be helpful and serve as indicator of how "typical" a pattern is compared to other patterns (Alasuutari, 1995; Silverman, 2005). Reporting of findings is based on the presentation of a large amount of data extracts on which the researcher has based his or her claims and which allows the reader to critically assess the analysis.

Analytical Induction

Identifying typical patterns requires that sufficient attention is also given to deviant or "negative" cases, i.e., occurrences which do not appear to fit the prototypical pattern and which contradict initial hypotheses. Grounding of the data also requires a recursive procedure, i.e. the researcher begins with formulating a tentative hypothesis about a pattern found in relation to the research question, and then proceeds by testing this framework through successive passes through the data and revises it where necessary (e.g.; Lincoln & Guba, 1985; Strauss, 1987). Analytic induction thus forces revisions of initial hypotheses that will make the analysis valid when applied to an increasingly diverse range of cases. Deviant or negative cases may, however, not necessarily show that a pattern is not normative, but may be the exception that proves the rule (Potter, 1996, 2007). Since analysis follows an inductive strategy and aims at developing an argument that accounts for all data, deviant case analysis does not serve to explain for the variance but are analyzed to see if the overall argument needs to be modified (Alasuutari, 1995).

An example for coherence between qualitative design and data presentation with respect to the last two criteria would for example be the followin: based on systematic comparison of all instances of the codes/categories, a reoccurring pattern is described. Deviant cases are discussed as revealing further insights to research question. An example of low coherence between a qualitative design and data presentation would be if findings are reported based on a frequency count of occurrence of codes (Smith, 2009).

Intersubjective Consensus

Intersubjective consensus refers to the degree to which the data and interpretations have been validated by others. Two forms of intersubjective consensus are discussed in the literature: 'Consensus validity' refers to the degree to which the data and interpretations have been presented to a research group in order to disclose one's own blind spots and to discuss working hypotheses and results with them. 'Communicative validation' refers to the communicative validation with members of the fields (either the participants themselves or members of a given socio-cultural group), and is therefore also referred to as 'member checks'. Member checks with the actual participants of the study, while considered to be very fruitful, for instance, in interview studies that aim at learning about people's subjective theories, are not considered useful or appropriate, however, in discourse analysis (DA) since it presumes that many repertoires and discursive strategies are not in the informants' awareness (Elliot et al., 1999; Silverman, 2001; Taylor, 2001a). DA, in contrast, offers the possibility of the *reader's evaluation*: materials are presented in a form that allows readers to make their own checks and judgments (e.g., Potter, 1996; Potter, 2007). Moreover, in CA and DA, *participants' orientation* is a common means to evaluate a possible interpretation of an utterance: for instance, when someone provides an 'acceptance' as response to another person's utterance it provides evidence that what came before was an 'invitation'.

Coherence

This criterion refers to the degree to which the interpretation is internally consistent, comprehensive, and persuasive, and the degree to which findings are coherent with previous studies. Inferences must be logical, plausible, and sufficiently grounded in the data.

Systematic Proceeding

This criterion refers to the degree to which analysis has been conducted in a systematic way and is based on accepted procedures for analysis and has correctly applied the relevant analytical steps of a specific procedure.

Reflection of a Researcher's Subjectivity

This criterion refers to the degree to which the researcher's subjectivity has been reflected in the analysis. For one, this includes the researcher's identity and his role as part of the social setting, as well as his or her influence during data assessment. It also includes whether the researcher's reflections, irritations, feelings etc. during analysis were documented (e.g. how the researcher's way of thinking and hence the interpretation changed in traversing the 'hermeneutic circle') which is considered to be an important source of information as they document how earlier inaccuracies in the interpretation were corrected (e.g., Mruck & Breuer, 2003, May; Stiles, 1993) and may be treated as data in their own right (e.g., Flick, 2009; Hammersley & Atkinson, 1995; Taylor, 2001a; Taylor, 2001b).

Conclusion and Outlook

Qualitative methods have gained increasing importance in the field of cross-cultural psychology. In order to ensure high standards of empirical research it is essential to be aware of how to ensure validity and to meet criteria of rigor. In this paper, I have tried to summarize what appears to be an emerging consensus of broad criteria suitable to evaluate qualitative research. Due to the heterogeneity of the field, it is very difficult to define criteria that will apply to all qualitative procedures. I have pointed out that the above criteria do not constitute a fixed set of standard criteria but need to be understood as broad criteria and that the researcher needs to further specify how these criteria apply to a particular research project and may need to adapt them accordingly. As a word of caution against a mechanistic understanding of these criteria I want to echo Willig (2008) by pointing out that "research methods are not recipes but ways of approaching questions, and the value of our research depends on the skill with which we manage to match our methods to our questions in the pursuit of knowledge and understanding" (Willig, 2008).

Criteria for evaluating qualitative research are meanwhile not only common part of many text books but also are increasingly integrated in the policies for reviewers, for instance within the the British Sociological Association – Medical Sociology (Seale, 1999), the National Institutes of Health, Office of Behavioral and Social Sciences – Public Health (Nih, 2001), in Clinical Psychology (Elliot, Fischer & Rennie, 1999) and are also presently discussed within the German Research Foundation (DFG). The British Journal of Occupational and Organizational Psychology even lays open its criteria for evaluating papers using qualitative research methods on its website: <http://www.bpsjournals.co.uk/journals/joop/qualitative-guidelines.cfm>.

These developments should encourage us to produce high quality qualitative cross-cultural research that will contribute to the further advancement of the field of cross-cultural psychology.

References

- Alasuutari, P. (1995). *Researching Culture. Qualitative Method and Cultural Studies*. London: Sage.
- Berry, J. W., Poortinga, Y. H. & Pandey, J. (Hrsg.) (1997). *Handbook of cross-cultural psychology, Vol. 1: Theory and method* (2. Aufl.). Needham Heights, MA US: Allyn & Bacon
- Boesch, E. E., & Straub, J. (2006). Kulturpsychologie. Prinzipien, Orientierungen, Konzeptionen. In G. Trommsdorff, & H. Kornadt (Eds.), *Kulturvergleichende Psychologie. Enzyklopädie der Psychologie*. Serie VII. Themenbereich C „Theorie und Forschung“ (pp. 25-95). Göttingen: Hogrefe.
- Bohnsack, R. (2001). Typenbildung, Generalisierung und Komparative Analyse: Grundprinzipien der Dokumentarischen Methode. In R. Bohnsack, I. Nentwig-Gesemann & A. Nohl (Eds.), *Die Dokumentarische Methode und ihre Forschungspraxis. Grundlagen Qualitativer Sozialforschung* (pp. 225-252). Opladen: Leske + Budrich.
- Camic, P. M., Rhodes, J. E., Yardley, L., Rhodes, J. E., & Yardley, L. (2003). *Qualitative research in psychology. Expanding perspectives in methodology and design*. Washington, DC: American Psychological Association .
- Corbin, J., & Strauss, A. L. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qualitative Sociology*, 13(1), 3-21.
- Coyle, A. (2000). Discourse analysis. In G. M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.), *Research methods in psychology* (2nd ed.). (pp. 251-268): Sage Publications Ltd.
- Drisko, W. J. (1997). Strengthening qualitative studies and reports. *Journal of Social Work Education*, 33(1), 185-197.
- Elliot, R., Fischer, C. T., & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38, 215-229.
- Flick, U. (2009). *An Introduction to Qualitative Research* (4th ed.). London: Sage.
- Flick, U. (2008). *Managing Quality in Qualitative Research* - Book 8 of the SAGE Qualitative Research Kit (ed. by U. Flick). London/ Thousand Oaks, CA/ Dehli: Sage
- Gaskell, G., & Bauer, M. W. (2000). Towards Public Accountability: beyond Sampling, Reliability and Validity. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative Researching with Text, Image and Sound: A practical Handbook* (pp. 336-350). London: SAGE Publications Ltd.
- Geertz, C. (1973). Thick Description: Toward an Interpretative Theory of Culture. In: *The Interpretation of Cultures*. New York: Basic Books.
- Glaser, B. G. (1978). *Theoretical sensitivity*. Mill Valley, CA: Sociology Press.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory. strategies for qualitative research*. Chicago: Aldine
- Guba, E. G., & Lincoln, Y. S. (1982). Epistemological and Methodological Bases of Naturalistic Inquiry. *Educational Communication and Technology*, 30(4), 233-252.
- Hammersley, M., & Atkinson, P. (1995). *Ethnography: Principles in Practice*. London and New York: Routledge.
- Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity.
- Jefferson, G. (1984). Transcription Notation. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Interaction: Studies in Conversation Analysis*. New York: Cambridge University Press.
- Kim, U., Park, Y., & Park, D. (2000). The challenge of cross-cultural psychology: The role of the indigenous psychologies. *Journal of Cross-Cultural Psychology*, 31(1), 63-75.
- Lamnek, S. (2006). *Qualitative Sozialforschung : Lehrbuch* 4., vollst. überarb. Aufl., [Nachdr.]. Weinheim: Beltz PVU.
- Lincoln, Y. S. (1995). Emerging criteria for quality in qualitative and interpretative research. *Qualitative Inquiry*, 1(3), 275-289.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. London: Sage.
- Lyons, E., & Coyle, A. (Eds.). (2007). *Analyzing qualitative data in psychology*. London: Sage.
- Mey, G. (2010). Qualitative developmental psychology. In A. Toomela, & J. Valsiner (Eds.), *Methodological thinking in psychology: 60 years gone astray? (advances in cultural psychology: Constructing human development)* (pp. 209-230) Information Age Pub Inc.
- Mruck, K., & Breuer, F. (2003, May). Subjectivity and Reflexivity in Qualitative Research—The FQS Issues [30 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal], 4(2), Art. 23 Available at: <http://www.qualitative-research.net/fqs-texte/1-00/1-00mruck-e.htm>, [Date of Access: March 17, 2008].
- Mey, G. & Mruck, K. (Eds.) (2010), *Handbuch Qualitative Forschung in der Psychologie*. Wiesbaden: VS Verlag
- Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods* (3rd edition). Thousand Oaks, CA: Sage Publications.
- Phillips, L., & Jorgensen, M. W. (2002). *Discourse Analysis as Theory and Method*. London: SAGE Publications Ltd.
- Potter, J. (1996). Discourse Analysis and Constructionist Approaches: Theoretical Background. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social science* (pp. 125-140). Leicester: BPS

Books.

- Potter, J. (2004). Discourse analysis as a way of analysing naturally occurring talk. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice* (2nd Ed.). London: Sage.
- Potter, J. (2007). Discourse Analysis and Discursive Psychology. In P. M. Camic, J. E. Rhodes & L. Yardley (Eds.), *Qualitative Research in Psychology* (pp. 73-94). Washington, DC: American Psychological Association.
- Ratner, C., & Hui, L. (2003). Theoretical and methodological problems in cross-cultural psychology. *Journal for the Theory of Social Behaviour*, 33(1), 67-94.
- Richardson, J. T. E. (1996). *Handbook of Qualitative Research Methods for Psychology and Social Science*. BPS Books: Leicester.
- Seale, C., & Silverman, D. (1997). Ensuring rigour in qualitative research. *European Journal of Public Health*, 7, 379-384.
- Silverman, D. (2001). *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction* (2 ed.). London: SAGE Publications Ltd.
- Silverman, D. (2005). *Doing Qualitative Research*. London: Sage.
- Smith, J. A. (Ed.). (2008). *Qualitative psychology. A practical guide to research methods* (2nd ed.). Los Angeles; London; New Delhi; Singapore: Sage.
- Steinke, I. (2004). Quality criteria in qualitative research. In Flick, U.v.K., Ernst & Steinke, Ines (Ed.), *A companion to qualitative research. [original: Qualitative Forschung. Ein Handbuch, 2004]* (pp. 184-190). London: Sage.
- Stiles, W. B. (1993). Quality control in qualitative research. *Clinical Psychology Review*, 13(6), 593-618.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists.*: Cambridge University Press.
- Strauss, A. L., & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. London: Sage.
- Taylor, S. (2001a). Evaluating and Applying Discourse Analytical Research. In M. Wetherell, S. Taylor & S. J. Yates (Eds.), *Discourse as Data. A Guide for Analysis*. (pp. 311-330). London: Sage.
- Taylor, S. (2001b). Locating and Conducting Discourse Analytical Research. In M. Wetherell, S. Taylor & S. J. Yates (Eds.), *Discourse as Data - A Guide for Analysis*. London: Sage
- Weber, M. (1968). Die ‚Objektivität‘ sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen
- Willig, C. (2008). *Introducing qualitative research in psychology: Adventures in theory and method* (2nd edition ed.). Berkshire, UK: Open University Press.