Selecting criteria to evaluate qualitative research ®

MARIA NORTHCOTE
Avondale College of Higher Education
Australia

Abstract

While the evaluation of quantitative research frequently depends on judgements based on the "holy trinity" of objectivity, reliability and validity (Spencer, Ritchie, Lewis, & Dillon, 2003, p. 59), applying these traditional criteria to qualitative research is not always a "good fit" (Schofield, 2002). Instead, educational researchers who engage in qualitative research have suggested various sets of alternative criteria including: transferability, generalisability, ontological authenticity, reciprocity, dependability, confirmability, reflexivity, fittingness, vitality and, even, sacredness and goodness (Creswell, 2002; Garman, 1996; Guba & Lincoln, 1989; Patton, 2002; Spencer et al., 2003; Stige, Malterud, & Midtgarden, 2009).

While over one hundred sets of qualitative research criteria have been identified (Stige et al., 2009), some researchers warn against the absolute application of any criteria to qualitative research which is, by its nature, wide-ranging and varied, and does not necessarily lend itself to the straightforward application of any evaluation criteria. Nevertheless, whether or not criteria are applied at all in the research evaluation process, postgraduate students face a number of decisions associated with the process of evaluating qualitative research: 1) whether or not to adopt a set of appraisal criteria; 2) which criteria to select, if criteria are used; and 3) how to apply alternative approaches to criteria-focused evaluation. These decisions often require a paradigm shift (Khun, 1962) in the way postgraduate students perceive and approach their research. The messiness and complexity associated with such decisions can be confronting. This paper examines a number of approaches used by researchers to evaluate qualitative investigations in educational research.

Introduction

> *" … it is important to regularly dialogue about what makes for good qualitative research" (Tracy, 2010, p. 837)*

The problem of how to evaluate the quality of qualitative research is not a new phenomenon (Lincoln & Guba, 2000). On an individual basis, many qualitative researchers face the day-to-day dilemma of determining how to evaluate the quality of their research, especially when the "holy trinity" of objectivity, reliability and validity (Spencer et al., 2003, p. 59) does not suffice. Qualitative research is at times unquestionably evaluated using the criteria of quantitative research (Cohen & Crabtree, 2008; Lather, 2004). In fact, Howe (2004) argues that the dominant paradigm in educational research is scientifically based experimental research. Assumptions about the superiority of quantitative research can diminish the significance and value of qualitative research by overlooking the importance of matching fit for purpose when it comes to research paradigm, the epistemological stance of the researcher and the selection of methodology. This is something researchers were cautioned about over a decade ago in Lincoln and Guba's account of *paradigmatic controversies, contradictions and emerging confluences* (2000).

Cohen and Crabtree (2008) found that reviewers of academic research papers and research grant applications are more likely to adopt generic criteria to evaluate research, rather than adopting criteria relevant to the specific qualitative research used. Furthermore, Alvesson and Skoldberg question the assumption that pure facts and data are the "solid bedrock of research" (2009, p. 3) and suggest that qualitative research requires, by its nature, creative and open minded approaches that acknowledge blurriness, complexity and subjectivity. When considering the origins of expectations about research is conducted and evaluated, Lather (2004) found the application of quantitative evaluation criteria to qualitative research often takes place systematically at authoritative and government levels. This established practice renders the overuse of quantitative criteria as a matter of course and can make it difficult for the lone researcher or even groups of researchers to combat or refute the inappropriate application of mismatched criteria, intended for quantitative research, to qualitative research.

The problems associated with how to evaluate the quality of qualitative research are encountered by experienced researchers in many disciplines so it comes as no surprise to note that postgraduate higher degree students have difficulty in deciding how to evaluate their own qualitative research. Consequently, supervisors of postgraduate students may find themselves with the responsibility of instructing their students about how to evaluate their own qualitative research and the research of others. Since positivism is frequently viewed as the preferred and accepted method of research (Cohen & Crabtree, 2008) in some contexts, the value of qualitative research may be questioned by novice researchers (Freshwater, Cahill, Walsh, & Muncey, 2010, p. 497). Accordingly, the supervisor's role in this educative process may involve dispelling their students' misconceptions about the very nature of research itself.

This paper provides a discussion of the problems and possible challenges associated with the judgement of a piece of qualitative research using inappropriate criteria that may be more suited to quantitative research. The paper does not disregard the value of quantitative research but argues that a step in the postgraduate research student's learning journey should include close examination and identification of the tenets of qualitative *and* quantitative research, in conjunction with an analysis of the criteria or methods used to evaluate the quality of more than one type of research. At the conclusion of this discussion, a number of implications for supervisors of postgraduate higher degree students will be outlined, with a particular focus on guiding the choice of the novice researcher in how to evaluate the quality of their own research and the research of others.

The apple-orange problem: Evaluation of research

> *Imagine a beautiful, perfectly ripe, shiny, red apple, recently plucked from a healthy mature tree growing in a well maintained orchard. Now, imagine volunteering this apple as a candidate for "The Best Orange in the World" competition. The apple would be judged by criteria quite unrelated to the apple's natural characteristics. The apple would be judged by inappropriate criteria such as: dimpled orange skin, capable of producing orange juice and other qualities of an excellent orange. The poor old misjudged apple, despite its excellence, would be found seriously lacking if placed in the "The Best Orange in the World" competition. No matter how "good" the apple was in such a competition, its entry is doomed to failure. The apple is not a good fit when judged by criteria used to evaluate a good orange.*

How can we judge a good apple by the criteria of a good orange? This question is almost absurd in its inappropriateness. So, why is it that qualitative research is regularly evaluated using

criteria that guide the recognition of excellence in quantitative research: objectivity, reliability and validity, referred to by Spencer et al. (2003, p. 59) refer as the "holy trinity"?

While some researchers have been critical of evaluating qualitative methodologies from the theoretical standpoint of positivism (Lather, 2004), others openly admit that they apply positivist criteria when judging the quality of qualitative research. For example, after acknowledging "a long tradition of using qualitative case studies" in business-to-business marketing research, Beverland and Lindgreen analysed a collection of these case studies and found that: "… from a positivist viewpoint, there has been a steady improvement in how authors addressed issues of research quality in published qualitative case studies" (2010, p. 56). This process is problematic as it represents an attempt to evaluate a set of research studies using criteria that are not necessarily relevant to the research itself. The criteria are not a good fit with the research. The crucial moment of recognising whether or not the application of quantitative criteria is appropriate when judging the goodness of qualitative research, needs to be informed by an understanding of the paradigmic tenets of varied forms of research, research that is theoretically founded on epistemological, philosophical and methodological diversity.

While many researchers are aware of the need to match research paradigm with research methodology, some are not (Cohen & Crabtree, 2008; Lather, 2004). Such misunderstandings may misinform postgraduate researchers who are in the early stages of their higher degree journey. Thus, postgraduate higher degree students frequently require guidance when they encounter applications of mismatched criteria to qualitative research. Without asking the question, "Is this an appropriate way to evaluate this research?", the postgraduate student may find themselves in a position where they are unable to defend the theoretical and methodological directions they follow in the design and implementation of their research project. Evaluating the quality or otherwise of other researchers' work (for example, during the process of conducting a literature review), is a crucial skill to be developed in the postgraduate student. Understanding the importance of goodness of fit in relation to how research studies are conducted provides the postgraduate student with a strong foundation by which to conduct a critical literature review and, subsequently, to use the findings of the literature review to design their own research study. Furthermore, when in the stages of designing their own qualitative research studies, postgraduate research students may encounter questions such as the following from well-intentioned colleagues, publication reviewers and possibly some faculty staff:

- The sample is not large enough.
- How can you remove the researcher's bias from that study?
- How will you keep track of all the variables in that study?
- Would the same results be found if you repeated this research?
- Do you have statistical data to uphold your claims?

In the face of such questions, the higher degree research student would be best served by allowing the findings of the literature review to inform the research question and choice of methodology (Borrego, Douglas, & Amelink, 2009) which are informed by theoretically appropriate research paradigms. In some cases, quantitative methods may be most appropriate to adopt while in other cases, the intention and focus of the research may be best suited to an interpretivistic research paradigm and the associated implementation of qualitative methods (Greene & Caracelli, 2003). Whatever the choices required by the postgraduate student, they should be informed by the research question (Borrego et al., 2009), the research intention and context (Creswell, 2002), the nature of the research (Johnson & Christensen, 2012), the theoretical and methodological framework from which the research emerges (Cohen &

Crabtree, 2008) and by acknowledging the perspectives of the research participants (Alvesson & Skoldberg, 2009).

However, remaining resolute about being informed by such noble intents is not necessarily an easy, quick or effortless process. While struggling to align the research intention and the research method, the postgraduate student may well enter into a state of what Kiley describes as "stuckness" (2009, p. 293), where doctoral students, for example, encounter challenging threshold concepts about research. Although this position may be essential in the process of learning about the nature of research, remaining in this state of "stuckness" too long can cause unnecessary frustration for the postgraduate student and the supervisor. The role of the postgraduate supervisor may well be to assist the student to become "unstuck". The process of becoming "unstuck" may be catalysed by the postgraduate student engaging in a discussion about the epistemological implications of various research paradigms (Pintrich, 2002) and how epistemology impacts on research (Schwandt, 2000); What type of knowledge is the researcher searching for or constructing? How is this knowledge gathered, constructed, co-constructed and analysed? What type of knowledge is shared with and reported to the academic community at the conclusion of the research? Epistemology and its variant stances are central to the tenets of any type of research (Howe, 2004). The apple-orange problem described earlier in this paper is indicative of an epistemological muddle. Knowledge of how research paradigms are informed by specific epistemological stances could help postgraduate research students avoid becoming a casualty of such a conflict.

The problems associated with selecting appropriate criteria by which to evaluate the goodness of qualitative research is, therefore, not just related to the research paradigm and intention, but also to the epistemological beliefs held by both researchers and research participants. In some cases, the epistemology of the audience for whom the research findings are intended may also influence the nature of the research. In addition to questions of research paradigm and epistemological belief, the assessment of the quality of qualitative research is also dependent upon the field of inquiry and disciplinary field.

The challenge of how to assess qualitative research is evident in many fields including psychology (Cassell & Symon, 2011), engineering (Borrego et al., 2009), health (Freshwater et al., 2010; Hannes, Lockwood, & Pearson, 2010) and industrial marketing (Beverland & Lindgreen, 2010). In some cases, particular sets of criteria are recommended for specific types of qualitative research such as case study research (Beverland & Lindgreen, 2010), personal construct research (Viney & Nagy, 2011) and research in the field of health (Cohen & Crabtree, 2008; Freshwater et al., 2010; Hannes et al., 2010). The discrepancies and debates that are associated with the choices of how to evaluate qualitative research are evident across many disciplines and the postgraduate student will no doubt encounter such philosophical and paradigmic conflicts as they traverse the literature about research methodology.

The concept of "research" is frequently portrayed using metaphors of white-coated scientists busily identifying variables in experiments or grappling with statistical reports. A simple Google Image search for the term "researcher" results in an instant collection of graphic metaphors that feature test tubes, laboratory scientists hunched over microscopes and cartoons about the weird and wonderful uses of statistics. The "subjects" of inquiry are often inanimate – including microscopes, plastic models of atoms, statistical graphs, numeric data and bubbling liquids. There is rarely an interviewer, an ethnographer or an observer of human behaviour in sight. *Research as experiment* is the pervading paradigm, as it is in much popular media: internet, daily television news, movies and newspapers. When conducting research or being involved in research as a researcher-participant, these scientific-type metaphors for research are not helpful to the postgraduate research who may still be coming to terms with the underlying theoretical

tenets and the subsequent practical applications of qualitative research. Concepts that pervade and form the basis of qualitative research, such as researcher-as-participant (Patton, 2002), reflexivity (Alvesson & Skoldberg, 2009) and multiple voicing (Gergen & Gergen, 2000), are not necessarily part of the novice researcher's foundational understanding of research. As such, the assumption that quantitative research is "real research" is sometimes a point of troublesome knowledge (Perkins, 2006) and could be described as a threshold concept (Kiley, 2009; Meyer & Land, 2003) for postgraduate research students.



Figure 1: Results of a Google Image search for the term "Researcher"

With the discrepancies associated with quantitative and qualitative research in mind, the juxtaposition of judging the quality of qualitative research with criteria more suited to quantitative research can hinder the postgraduate researcher's progress. The following section of this paper aims to provide a set of associated concepts with which to examine, question and evaluate criteria that may be adopted in interpretivistic or qualitative research studies.

To use or not to use criteria?

Qualitative researchers are not known for their propensity to follow a pre-defined set of recipe-style guidelines. Nor are they particularly comfortable with applying a rigid set of standards to guide the selection or use of research methodologies. Rather, they attempt to recognise the diversity and complexity of their research participants and the context, and, subsequently, work *with* rather than *within* the boundaries or contexts of their research settings. Unlike the general consensus evident amongst researchers about the criteria by which to judge quantitative research methodologies, there is understandably less clarity and formulaic direction about which criteria to apply to qualitative research. Garman (1996) suggests that, unlike quantitative research, "qualitative research is relatively lacking in canons and conventions" (p. 5). Tong, Sainsbury and Craig (2007, p. 350) go so far as to suggest that "no consolidated reporting framework exists for any type of qualitative design". Because of the wide-ranging nature of

qualitative research, it is not necessarily a "unified field" (Cohen & Crabtree, 2008, p. 338) so the choice of which criteria to use, apart from the consideration of whether or not to use criteria at all, can be a dilemma for the experienced researcher and the postgraduate student alike.

Despite the confusion about how to evaluate qualitative research, the need exists to tackle this topic, especially for the postgraduate neophyte researcher. Spurred on by the pioneering work in this area by Guba and Lincoln (Guba & Lincoln, 1989; Lincoln, 1995; Lincoln & Guba, 1985) on *fourth generation criteria* and *emerging criteria for quality in qualitative*, many other researchers followed their lead by constructing sets of criteria by which to evaluate the goodness of qualitative and interpretive research (Creswell, 2002; Driessen, Van Der Vleuten, Schuwirth, Van Tartwijk, & Vermunt, 2005; Flick, 2004; Freebody, 2003; Patton, 2002; Tong et al., 2007; Viney & Nagy, 2011). Back in 2001, Geelan suggested that, "If positivist standards of validity and reliability are no longer considered appropriate for some forms of educational research, however, new standards for justification and representation, explicitly stated within the research, will be necessary" (p. 129).

The investigation of complex phenomena is not a straightforward process (Tong et al., 2007). While one group of researchers expound the value of using a new and relevant set of criteria for judging the quality of qualitative researchers, other researchers caution against the absolute application of any criteria or standards to qualitative research. Howe issues a general warning against a standards approach in the field of education: "the standards-setting movement in education is inherently about power and control" (Howe, 2004, p. 57).

Nevertheless, whether or not the research is "good research" is still an important question to ask, despite the research paradigm under which the research is conducted or experienced. Those who access findings from any type of research are first and foremost interested in the "goodness" of the research as such qualities reflect on the trustworthiness (Lincoln, 1995; Patton, 2002) of its findings. When making evaluative judgements about qualitative research, there are many sets of criteria that can be used to guide our judgement. Despite the varied positions proposed by the qualitative researchers themselves, postgraduate higher degree students will likely be faced with decisions about which, if any, set of criteria to apply to their own studies. Following is an account of some of these sets of criteria.

While not suggesting a pre-defined set of criteria to apply prescriptively to qualitative research, Viney and Nagy (2011) guardedly suggest "A set of *possible* criteria" that may be used in the particular research context of personal construct psychology. Their cautious use of the term *possible* represents a position where they have inched away from the position of an absolute ban in using criteria in qualitative research. Instead, general descriptive terms are often offered to explain the features of such criteria. For example, Guba and Lincoln suggest various types of authenticity (Guba & Lincoln, 1989), goodness (Lincoln & Guba, 2000), confirmability (Lincoln & Guba, 1985) and fairness (Lincoln & Guba, 2000). Flick (1992, 2004) uses the criteria of rigour, breadth and depth; Eisner (1991) refers to coherence, consensus and instrumental utility; and Mays and Pope (2000) simply adopt the principles of validity and relevance.

When grappling with how to evaluate portfolios, the content of which are typically more qualitative than qualitative, Driessen et al. (2005) used a number of research strategies (including member checking, prolonged engagement, triangulation) to ensure they achieved credibility and dependability of their evaluation processes. In the absence of a comprehensive checklist of criteria for reporting qualitative studies, Tong et al. (2007) created their own checklist including 32 items, which they refer to as a *consolidated criteria for reporting qualitative research (COREQ)*. In the context of nursing research, Freshwater et al. (2010) focus on the criteria of rigour and relevance, while Hannes, Lockwood and Pearson (2010) focus mainly on validity.

Cohen and Crabtree (2008), after analysing research studies that focused on rigour in qualitative research, found seven criteria for good qualitative research: (1) carrying out ethical research; (2) importance of the research; (3) clarity and coherence of the research report; (4) use of appropriate and rigorous methods; (5) importance of reflexivity or attending to researcher bias; (6) importance of establishing validity or credibility; and (7) importance of verification or reliability. This list-style set of criteria is supplemented by an emphasis on the significance of relevant language to analyse qualitative research. While acknowledging the vast diversity of qualitative research, Tracy (2010) recommends the adoption of a common language by which to discuss the quality of qualitative research. She contributes to the conversation about how to evaluate qualitative research by providing a set of eight key markers of quality in qualitative research: (a) worthy topic, (b) rich rigor, (c) sincerity, (d) credibility, (e) resonance, (f) significant contribution, (g) ethics, and (h) meaningful coherence".

Richardson's (1997) notion of the crystallisation of data is enabled through the analysis of both qualitative and quantitative data to ensure what Gergen and Gergen refer to as "multiple voicing" (2000, p. 1025). Spencer et al.'s (2003) framework for evaluating interpretivistic research proposes that good qualitative research should be contributory, defensible, rigorous and credible. They purport that the epistemological underpinnings of the research should also be acknowledged and featured in the research methodology and reporting method.

Other criteria, especially those suggested by Garman (1994, 1996), Richardson (1997) and Spencer et al. (2003), are largely associated with the aesthetics of research processes and products rather than just the rationality of inquiry. Such criteria convey a sense of the emotional intensity, vitality and even the spiritual or moral value of research processes and findings. A movement to incorporate these elements of research was predicted by Lincoln and Guba over ten years ago: "We may also be entering an age of greater spirituality within research efforts" (2000, p. 185). Freebody also acknowledged the aspect of research that was not preoccupied with rationalism by describing qualitative research as a "moral effort for researchers and readers" (Freebody, 2003, p. 218).

This recognition of the emotional, personal, moral and spiritual characteristics of research represents a diversion from the direct application of criteria by which to judge the quality of research, whether it be quantitative or qualitative. The process of acknowledging the aesthetics of research is associated with the researchers' and the participants' emotional involvement, the researcher's commitment and enthusiasm, and the manner in which the findings are presented to readers, especially in regards to how the research report or thesis communicates discovery, excitement and insight. The view of research that acknowledges logic, future impact and credibility, as well as emotional and spiritual pursuits, can be likened to Schulman's manner of defining scholarship as "acts of the mind or spirit that are undertaken in disciplined ways and subsequently made public so that members of one's intellectual community can judge their worth and then use them to support the more general program of the community" (1999, p. 160).

The bounds of this paper do not extend far enough to contain a full analytic or aesthetic account of the wide array of criteria, language, dimensions and questions that have been offered by researchers over the years by which to evaluate the quality of qualitative research. However, the following table attempts to provide an interconnected overview of some of these attempts. The criteria in the following table are presented in the form of guiding principles where were informed by the four central principles of Spencer et al.'s (2003) framework for evaluating interpretivistic research (contributory, rigorous, defensible, credible), as well as being informed by Garman's (1994, 1996) recommendations that research should be judged by its vitality and aesthetics. These principles have been transformed into guiding questions to assist the novice researcher to reflect on how these principles apply to the processes of designing and conducting qualitative research.

Table 1: Guiding Principles and Criteria used to Evaluate the Qualitative Nature and Methods of Research

Guiding Principle	Guiding Question	Specific Criteria from Literature*
Contributory in advancing wider knowledge or understanding about policy, practice, theory or a particular substantive field.	Have the findings of this study contributed to our knowledge and understanding of the educational beliefs of university teachers and students? How has being involved in the research benefited the participants?	Significant contribution (Tracy, 2010) Instrumental utility (Eisner, 1991); utility (Garman, 1994, 1996); useability Transferability or applicability; naturalistic generalisability (Stake, 1978; Stake & Trumbull, 1982); extrapolation (Patton, 1990, 2002); retrospective generalisability (Eisner, 1991); fit between the situation studied and others (Schofield, 2002); illuminative fertility (Shank & Villella, 2004) Future focus and contribution to research directions Ontological and educative authenticity (Guba & Lincoln, 1989) Reciprocity (Patton, 1990, 2002) Importance of the research (Cohen & Crabtree, 2008); worthy topic (Tracy, 2010)
Rigorous in conduct through the systematic and transparent collection, analysis and interpretation of qualitative data.	Are the methods used to gather, analyse, interpret and present the data rigorous, systematic and transparent? - Gathering data - Interpreting and analysing data - Reporting the findings	Openness and clarity (Cohen & Crabtree, 2008) Ethics; carrying out ethical research (Cohen & Crabtree, 2008; Tracy, 2010) Referential adequacy Resonance (Tracy, 2010) Use of multiple sources, multiple voicing (Gergen & Gergen, 2000); fairness (Lincoln & Guba, 2000) Rich rigour (Tracy, 2010); rigour (Cohen & Crabtree, 2008; Flick, 2004; Freshwater et al., 2010; Viney & Nagy, 2011) Transparency of data gathering methods, prolonged engagement (Driessen et al., 2005) Thick description Systematic primary and secondary interpretations Transparency of data analysis methods Empathic neutrality (Patton, 1990, 2002) Verite' (Garman, 1994, 1996); fittingness Descriptive validity (Maxwell, 2002); verisimilitude (Garman, 1994, 1996); investigative depth (Shank & Villella, 2004) Interpretive validity (Maxwell, 2002); interpretive adequacy (Shank & Villella, 2004) Credibility (Driessen et al., 2005; Tracy, 2010; Viney & Nagy, 2011)

Table 1: Guiding Principles and Criteria used to Evaluate the Qualitative Nature and Methods of Research

Guiding Principle	Guiding Question	Specific Criteria from Literature*
Defensible in design by providing a research strategy that can address the evaluative questions posed.	Is the research design of this study defensible and trustworthy, and linked to the study's research questions?	Goodness (Lincoln & Guba, 2000; Smith, 1993); integrity (Garman, 1994, 1996), fittingness or consistency Dependability Auditability; audit trail (Whitt, 1991); transparency of methods, participatory accountability (Shank & Villella, 2004) Confirmability (Lincoln & Guba, 1985) Reflexivity (Alvesson & Skoldberg, 2009; Cohen & Crabtree, 2008; Creswell, 2002); transparency of researcher's ideas, or external reliability (LeCompte & Goetz, 1982) Triangulation (Driessen et al., 2005); crystallisation (Richardson, 1997)
Credible in claim through offering well-founded and plausible arguments about the significance of the evidence generated.	Are the findings credible and supported by evidence?	Peer debriefing, peer confirmability, member checking (Driessen et al., 2005) Participant debriefing, participant confirmability Consensus and meaningful coherence (Cohen & Crabtree, 2008; Eisner, 1991; Tracy, 2010) Theoretical, descriptive validity and interpretive (Maxwell, 2002); relevance (Freshwater et al., 2010; Mays & Pope, 2000) Verisimilitude (Garman, 1994, 1996); investigative depth (Shank & Villella, 2004) Interpretive adequacy (Shank & Villella, 2004)
Affective in nature by acknowledging the excitement associated with research discoveries, the emotional involvement of the participants and the enthusiasm of the researcher.	Do the research processes and findings communicate the emotional elements of how the participants and the researcher engaged in the research study?	Vitality and aesthetics (Garman, 1994, 1996) Sacredness (Lincoln & Guba, 2000); sacred places (Richardson, 1997) Sincerity (Tracy, 2010) Moral effort for researchers and readers (Freebody, 2003)

While Table 1 demonstrates the various terms and phrases associated with many researchers' endeavours to create a set of criteria appropriate to the process of evaluating qualitative research, the table's content also illustrates the complexity of selecting the most appropriate criteria for a particular research study, if any are appropriate at all. In fact, the criteria used to evaluate qualitative studies appear to be under just as much scrutiny as the research itself.

After considering the positions put forward in this paper regarding the challenge of how to evaluate the quality of qualitative research, the postgraduate student may be left with a different question from "What criteria should I use to judge the quality of my qualitative research?" After an analysis of the paradigmic diversity of the theoretical influences and foundations of qualitative research, the higher degree student may, understandably, approach

their supervisor with the question, “When designing, conducting and experiencing qualitative research, should I use evaluation criteria at all?”

The postgraduate supervisor’s challenge

With the aforementioned dilemmas and complexities associated with evaluating qualitative research in mind, supervisors of postgraduate research students are often left with more than one challenge associated with the process of guiding their students through the mire of choices that present themselves when conducting qualitative research. Following are a set of suggestions that may, when considered alongside the student’s personal characteristics and needs, the nature of the issue being researched and the supervisor’s disposition, provide some useful recommendations for how to facilitate the learning of postgraduate students about the theoretical and practical nature of qualitative research and how to fairly evaluate it.

Epistemology. The nature of research is irretrievably connected to the nature of knowledge and epistemological beliefs (Howe, 2004; Pintrich, 2002; Schwandt, 2000). Without engaging in a reflective and ongoing discussion about how research methodologies and strategies are connected to and driven by the researcher’s and the research participants’ attitudes to knowledge and truth, the postgraduate student may struggle with making theoretically informed decisions about how to design and conduct a practical qualitative inquiry. Greene and Caracelli (2003), in their chapter titled, *Making paradigmic sense of mixed methods practice*, advocate the recognition and analysis of philosophical beliefs in conjunction with the practice of inquiry and research. Before choosing an appropriate research paradigm to operate within, and definitely before any methodological choices are made, the postgraduate student could be encouraged to explore the nature of epistemology in research, their own epistemological beliefs (Hofer & Pintrich, 2002) and the beliefs of others, and make plans for how to incorporate conflicting, blurred and varied epistemologies ranging from naïve to sophisticated (Brownlee, 2002) into their research.

White-coat research perceptions. While some higher degree research students may enrol in their postgraduate course of study with an understanding of the diversity of qualitative, quantitative and mixed mode research, others may bring to their study a narrower view of research, possibly influenced by the prevailing view of society about traditional research (Alvesson & Skoldberg, 2009; Cohen & Crabtree, 2008; Lather, 2004). The supervisor of such students may find it helpful to facilitate the expansion of their students’ perspectives about the nature and value of varied forms of research. Assisting their students to understand research as a process which can extend beyond the bounds of experimental methods, the postgraduate supervisor may play an instrumental role in preventing the philosophical narrowness associated with what Lincoln and Cannella speak of as “methodological conservatism” (2004, p. 10). Encouraging postgraduate student researchers to focus on the research question (Borrego et al., 2009, p. 53) prior to the choice of methodology may also assist in achieving the aim to broaden and deepen the postgraduate student’s understanding of the nature of research.

Thesis examiners. Although this decision is not necessarily left in the postgraduate student’s hands, a choice of appropriate readers of the student’s final thesis should be seriously considered by the student’s supervisor to ensure that the reader’s philosophical, epistemological and theoretical position is commensurate with the nature of the research represented by the final thesis. The theoretical disposition of the thesis reader should be closely analysed through an exploration of their previous work. Failure to match the research in the thesis with an appropriate reader is akin to entering an apple into a competition for the best orange, mentioned earlier in this paper.

The question of criteria. Whether or not to use criteria and, if appropriate, whether to use criteria to evaluate the postgraduate student's research is a question that should be asked. The answer to this question should ideally be considered through conversations between the supervisor and the postgraduate student, in which epistemological positions are addressed alongside decisions about practical methodological issues and appropriate research paradigms. An analysis of the nature of the research being conducted could be the driving force behind how or if evaluative criteria are used by the postgraduate researcher.

Conclusion

Although this paper does not claim to offer a *magic bullet* for solving problems about how, why or whether to evaluate the "goodness" of qualitative research, it provides a platform from which higher degree students and their supervisors may consider the messy trails they find themselves traversing in the fields of qualitative research. The paper offers a collection of issues to consider, a set of questions to ask and multiple sets of criteria to explore when attempting to determine the quality of qualitative research. While posing the question of "What evaluative criteria are most appropriate for judging the goodness of qualitative research?", the paper also asks the question "Should criteria be used at all?" Rather than attempting to squeeze the round peg of qualitative research into the square hole that is the home of much qualitative research, this paper highlights the intrinsic value of qualitative research and the ensuing recognition of the need to judge such research by the very tenets of its being. Instead of comparing apples with oranges, or comparing qualitative research with quantitative research, the paper has argued from the premise of soundly and fairly judging the quality of apples based on the qualities of good apples, of judging the quality of oranges based on the qualities of good oranges, and of judging the quality of qualitative research based on the qualities of good qualitative research.

References

Alvesson, M., & Skoldberg, K. (2009). Reflexive methodology: New vistas for qualitative research (2nd ed.). London: Sage.

Beverland, M., & Lindgreen, A. (2010). What makes a good case study? A positivist review of qualitative case research published in Industrial Marketing Management, 1971–2006. Industrial Marketing Management, 39(1), 56–63.

Borrego, M., Douglas, E. P., & Amelink, C. T. (2009). Quantitative, qualitative, and mixed research methods in engineering education. Journal of Engineering Education, 98(1), 53-66.

Brownlee, J. (2002, 7-10 July). Students learning to teach: Conversing with students about their epistemological beliefs. Paper presented at the Quality Conversations: Higher Education Research and Development Society of Australasia (HERDSA) Conference, Edith Cowan University, Perth, Western Australia.

Cassell, C., & Symon, G. (2011). Assessing 'good' qualitative research in the work psychology field: A narrative analysis. Journal of Occupational and Organizational Psychology, 84(4), 633–650.

Cohen, D. J., & Crabtree, B. F. (2008). Evaluative criteria for qualitative research in health care: Controversies and recommendations. Annals of Family Medicine, 6(4), 331-339.

Creswell, J. (2002). Educational research: Planning, conducting and evaluating quantitative and qualitative research. New Jersey: Pearson Education.

Driessen, E., Van Der Vleuten, C., Schuwirth, L., Van Tartwijk, J., & Vermunt, J. (2005). The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: A case study. Medical Education, 39(2), 214–220.

Lather, P. (2004). This is your father's paradigm: Government intrusion and the case of qualitative research in education. Qualitative Inquiry, 10(1), 15-34.

LeCompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. Review of Educational Research, 52(1), 31-60.

Lincoln, Y. S. (1995). Emerging criteria for quality in qualitative and interpretive research. Qualitative Inquiry, 1(3), 275-289.

Lincoln, Y. S., & Cannella, G. S. (2004). Dangerous discourses: Methodological conservatism and governmental regimes of truth. Qualitative Inquiry, 10(1), 5-14.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. California: Sage Publications.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research (2nd ed., pp. 163-188). California: Sage Publications.

Maxwell, J. A. (2002). Understanding and validity in qualitative research. In A. M. Huberman & M. B. Miles (Eds.), The qualitative researcher's companion (pp. 37-64). Thousand Oaks & London: Sage Publications.

Mays, N., & Pope, C. (2000). Quality in qualitative research. In C. Pope & N. Mays (Eds.), Qualitative research in healthcare. London: BMJ.

Meyer, J., & Land, R. (2003). Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the

Eisner, E. W. (1991). The enlightened eye: Qualitative inquiry and the enhancement of educational practice. New York: Macmillan Publishing Company.

Flick, U. (1992). Triangulation revisited: Strategy of validation or alternative? Journal for the Theory of Social Behaviour, 22(2), 175-197.

Flick, U. (2004). Design and process in qualitative research (B. Jenner, Trans.). In U. Flick, E. v. Kardoff & I. Steinke (Eds.), A companion to qualitative research (pp. 146-152). London: Sage Publications.

Freebody, P. (2003). Qualitative research in education: Interaction and practice. London: Sage.

Freshwater, D., Cahill, J., Walsh, E., & Muncey, T. (2010). Qualitative research as evidence: Criteria for rigour and relevance. Journal of Research in Nursing, 15(6), 497-508.

Garman, N. (1994, 5 August). Qualitative inquiry: Meaning and menace for educational researchers (Keynote address). Paper presented at the Mini-Conference: Qualitative Approaches in Educational Research, The Flinders University of South Australia.

Garman, N. (1996). Qualitative inquiry: Meaning and menace for educational researchers. In P. Willis & B. Neville (Eds.), Qualitative research practice in adult education (pp. 11-29). Ringwood, Victoria: David Lovell Publishing.

Geelan, D. R. (2001). Feyerabend revisited: Epistemological anarchy and disciplined eclecticism in educational research. Australian Educational Researcher, 28(1), 129-146.

Gergen, M. M., & Gergen, K. J. (2000). Qualitative inquiry. In Y. S. Lincoln & E. G. Guba (Eds.), Handbook of qualitative research (2nd ed., pp. 1025-1046). California: Sage Publications.

Greene, J. C., & Caracelli, V. J. (2003). Making paradigmic sense of mixed methods practice. In A. Tashakkori & C. Teddlie (Eds.), Handbook of mixed methods in social & behavioural research (pp. 91-110). Thousand Oaks, CA: Sage.

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

Hannes, K., Lockwood, C., & Pearson, A. (2010). A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. Qualitative Health Research, 20(12), 1736-1743.

Hofer, B. K., & Pintrich, P. R. (Eds.). (2002). Personal epistemology: The psychology of beliefs about knowledge and knowing. Mahwah, N.J.: Lawrence Erlbaum Associates.

Howe, K. R. (2004). A critique of experimentalism. Qualitative Inquiry, 10(1), 42-61.

Johnson, B., & Christensen, L. (2012). Educational research: Quantitative, qualitative, and mixed approaches (4th ed.). Thousand Oaks, CA: Sage.

Khun, T. S. (1962). The structure of scientific revolutions. Chicago: University of Chicago Press.

Kiley, M. (2009). Identifying threshold concepts and proposing strategies to support doctoral candidates. Innovations in Education and Teaching International, 46(3), 293-304.

disciplines. Edinburgh: University of Edinburgh.

Patton, M. Q. (1990). Qualitative evaluation and research methods. Newbury Park, CA: Sage Publications.

Patton, M. Q. (2002). Qualitative research and evaluation methods (3rd ed.). Thousand Oaks, California: Sage Publications.

Perkins, D. (2006). Constructivism and troublesome knowledge. In J. Meyer & R. Land (Eds.), Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge (pp. 33-47). New York: Routledge.

Pintrich, P. R. (2002). Future challenges and directions for theory and research on personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 389-414). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Richardson, L. (1997). Fields of play: Constructing an academic life. New Brunswick, New Jersey: Rutgers University Press.

Schofield, J. W. (2002). Increasing the generalisability of qualitative research. In A. M. Huberman & M. B. Miles (Eds.), The qualitative researcher's companion (pp. 171-203). Thousand Oaks & London: Sage Publications.

Schulman, L. (1999). Professing educational scholarship. In E. C. Lagemann & L. S. Shulman (Eds.), Issues in education research: Problems and possibilities (pp. 159-165). San Francisco: Jossey-Bass Publishers.

Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry. In Y. S. Lincoln & E. G. Guba (Eds.), Handbook of qualitative research (2nd ed., pp. 189-213). California: Sage Publications.

Shank, G., & Villella, O. (2004). Building on new foundations: Core principles and new directions for qualitative research. The Journal of Educational Research, 98(1), 46-55.

Smith, J. (1993). After the demise of empiricism: The problem of judging social and educational inquiry. Norwoods, New Jersey: Ablex.

Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). Quality in qualitative evaluation: A framework for assessing research evidence. London: National Centre for Social Research, Government Chief Social Researcher's Office, UK.

Stake, R. E. (1978). The case-study method in social inquiry. Educational Researcher, 7, 5-8.

Stake, R. E., & Trumbull, D. J. (1982). Naturalistic generalizations. Review Journal of Philosophy and Social Science, 7, 1-12.

Stige, B., Malterud, K., & Midtgarden, T. (2009). Toward an agenda for evaluation of qualitative research. Qualitative Health Research, 19(10), 1504-1516.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. International Journal of Qualitative Health Care 19(6), 349-357.

Tracy, S. J. (2010). Qualitative quality: Eight "Big-Tent" criteria for excellent qualitative research. Qualitative Inquiry, 16(10), 837-851.

Viney, L. L., & Nagy, S. (2011). Qualitative methods in personal construct research: A set of possible criteria. In P. Caputi, N. Crittenden, L. L. Viney & B. M. Walker (Eds.), Personal Construct Methodology: John Wiley & Sons.

Whitt, E. J. (1991). Artful science: A primer on qualitative research methods. Journal of College Student Development, 32, 406–415.

Corresponding author
Maria Northcote
maria.northcote@avondale.edu.au