Strategies for ensuring trustworthiness in qualitative research projects

Andrew K. Shenton
Division of Information and Communication Studies, School of Informatics, Lipman Building, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK

Received 14 November 2003
Accepted 6 January 2004

Although many critics are reluctant to accept the trustworthiness of qualitative research, frameworks for ensuring rigour in this form of work have been in existence for many years. Guba’s constructs, in particular, have won considerable favour and form the focus of this paper. Here researchers seek to satisfy four criteria. In addressing credibility, investigators attempt to demonstrate that a true picture of the phenomenon under scrutiny is being presented. To allow transferability, they provide sufficient detail of the context of the fieldwork for a reader to be able to decide whether the prevailing environment is similar to another situation with which he or she is familiar and whether the findings can justifiably be applied to the other setting. The meeting of the dependability criterion is difficult in qualitative work, although researchers should at least strive to enable a future investigator to repeat the study. Finally, to achieve confirmability, researchers must take steps to demonstrate that findings emerge from the data and not their own predispositions. The paper concludes by suggesting that it is the responsibility of research methods teachers to ensure that this or a comparable model for ensuring trustworthiness is followed by students undertaking a qualitative inquiry.

Keywords: Qualitative methods, research

1. Introduction

The trustworthiness of qualitative research generally is often questioned by positivists, perhaps because their concepts of validity and reliability cannot be addressed in the same way in naturalistic work. Nevertheless, several writers on research methods, notably Silverman [1], have demonstrated how qualitative researchers can incorporate measures that deal with these issues, and investigators such as Pitts [2] have attempted to respond directly to the issues of validity and reliability in their own qualitative studies. Many naturalistic investigators have, however, preferred to use different terminology to distance themselves from the positivist paradigm. One such author is Guba, who proposes four criteria that he believes should be considered
by qualitative researchers in pursuit of a trustworthy study [3]. By addressing similar issues, Guba’s constructs correspond to the criteria employed by the positivist investigator:

a) credibility (in preference to internal validity);
b) transferability (in preference to external validity/generalisability);
c) dependability (in preference to reliability);
d) confirmability (in preference to objectivity).

Although as recently as the mid 1990s Lincoln wrote that the whole area of qualitative inquiry was “still emerging and being defined” [4], Guba’s constructs have been accepted by many. This paper considers the criteria in detail and suggests provisions that the qualitative researcher may employ to meet them. The strategies advocated are based on the experience gained by Shenton when undertaking a qualitative PhD study devoted to the information-seeking behaviour of school-aged children [5].

2. Credibility

One of the key criteria addressed by positivist researchers is that of internal validity, in which they seek to ensure that their study measures or tests what is actually intended. According to Merriam, the qualitative investigator’s equivalent concept, i.e. credibility, deals with the question, “How congruent are the findings with reality?” [6] Lincoln and Guba argue that ensuring credibility is one of most important factors in establishing trustworthiness [7]. The following provisions may be made by researchers to promote confidence that they have accurately recorded the phenomena under scrutiny:

a) the adoption of research methods well established both in qualitative investigation in general and in information science in particular. Yin recognises the importance of incorporating “correct operational measures for the concepts being studied” [8]. Thus, the specific procedures employed, such as the line of questioning pursued in the data gathering sessions and the methods of data analysis, should be derived, where possible, from those that have been successfully utilised in previous comparable projects. In terms of investigation of information-seeking behaviour, the work of Dervin has proved particularly influential in this regard. In their study of the information needs of Seattle’s residents, Dervin et al. initially invited participants to reflect on situations “where you needed help... where you didn’t understand something... where you needed to decide what to do... or, where you were worried about something” [9]. Dervin’s respondents then described in detail a particular instance within one of these categories. Similar strategies have been used subsequently by Chen and Hernon [10], Poston-Anderson and Edwards [11] and Shenton [12] amongst others;
b) the development of an early familiarity with the culture of participating organisations before the first data collection dialogues take place. This may be achieved via consultation of appropriate documents and preliminary visits to the organisations themselves. Lincoln and Guba [13] and Erlandson et al. [14] are among the many who recommend “prolonged engagement” between the investigator and the participants in order both for the former to gain an adequate understanding of an organisation and to establish a relationship of trust between the parties. The danger emerges, however, that if too many demands are made on staff, gatekeepers responsible for allowing the researcher access to the organisation may be deterred from cooperating. The investigator may also react with some suspicion to the notion of prolonged engagement in view of the undesirable side effects that have been noted by Lincoln and Guba [15] and Silverman [16]. The former draw particular attention to the way in which investigators may become so immersed in the culture under scrutiny that their professional judgements are influenced;

c) random sampling of individuals to serve as informants. Although much qualitative research involves the use of purposive sampling, a random approach may negate charges of researcher bias in the selection of participants. As Preece notes, random sampling also helps to ensure that any “unknown influences” are distributed evenly within the sample [17]. Furthermore, it may be that a random method is particularly appropriate to the nature of the investigation. The work may, for example, take the form of a “collective case study” of the type described by Stake, in that multiple voices, exhibiting characteristics of similarity, dissimilarity, redundancy and variety, are sought in order to gain greater knowledge of a wider group [18], such as a more general population, rather than simply the individual informants who are contributing data. This form of research is also recognised by Hamel, Dufour and Fortin, who dub it a “macroscopic” case study, and emphasise the importance of appropriate selection tactics if the investigator is to be confident that informants are typical of members of a broader, “selected society” [19]. According to Bouma and Atkinson, “A random sampling procedure provides the greatest assurance that those selected are a representative sample of the larger group” [20]. A significant disadvantage of random method, however, stems from the fact that, since the researcher has no control over the choice of informants, it is possible that quiet, uncooperative or inarticulate individuals may be selected;

d) triangulation. Triangulation may involve the use of different methods, especially observation, focus groups and individual interviews, which form the major data collection strategies for much qualitative research. Whilst focus groups and individual interviews suffer from some common methodological shortcomings since both are interviews of a kind, their distinct characteristics also result in individual strengths. According to Guba [21] and Brewer and Hunter [22], the use of different methods in concert compensates for their individual limitations and exploits their respective benefits. Where possible,
supporting data may be obtained from documents to provide a background to and help explain the attitudes and behaviour of those in the group under scrutiny, as well as to verify particular details that participants have supplied. Opportunities should also be seized to examine any documents referred to by informants during the actual interviews or focus groups where these can shed more light on the behaviour of the people in question.

Another form of triangulation may involve the use of a wide range of informants. This is one way of triangulating via data sources. Here individual viewpoints and experiences can be verified against others and, ultimately, a rich picture of the attitudes, needs or behaviour of those under scrutiny may be constructed based on the contributions of a range of people. Van Maanen urges the exploitation of opportunities “to check out bits of information across informants” [23]. Such corroboration may, for example, take the form of comparing the needs and information-seeking action described by one individual with those of others in a comparable position. In addition, the investigator may draw informants from both users of an information service and the professionals who deliver it. Even in a user study, where the thrust of the work is likely to lie in analysing the ideas and experiences of users themselves, data provided by those responsible for the management and delivery of the service under scrutiny may well prove invaluable in order to check that supplied by the users, to help explain their attitudes and behaviour and to enhance the contextual data relating to the fieldwork site(s). Just as triangulation via data sources can involve the use of a diversity of informants, a range of documents may also be employed as source material. For example, documents created corporately by each participating organisation may be examined, as well as those relating to the organisation but produced externally. Further data dealing with the wider context in which the organisation is operating may be elicited from official publications.

Where appropriate, site triangulation may be achieved by the participation of informants within several organisations so as to reduce the effect on the study of particular local factors peculiar to one institution. Where similar results emerge at different sites, findings may have greater credibility in the eyes of the reader. The sampling of a range of people in different organisations may be employed to provide the diversity that underpins Dervin’s concept of “circling reality”, which she defines as “the necessity of obtaining a variety of perspectives in order to get a better, more stable view of ‘reality’ based on a wide spectrum of observations from a wide base of points in time-space” [24];

e) tactics to help ensure honesty in informants when contributing data. In particular, each person who is approached should be given opportunities to refuse to participate in the project so as to ensure that the data collection sessions involve only those who are genuinely willing to take part and prepared to offer data freely. Participants should be encouraged to be frank from the outset of each session, with the researcher aiming to establish a rapport in the opening
moments and indicating that there are no right answers to the questions that will be asked. Where appropriate, the independent status of the researcher should also be emphasised. Participants can, therefore, contribute ideas and talk of their experiences without fear of losing credibility in the eyes of managers of the organisation. It should be made clear to participants that they have the right to withdraw from the study at any point, and they should not even be required to disclose an explanation to the investigator. In many instances, such an unconditional right for subjects to withdraw may be a requirement that must be accepted by the researcher when seeking approval for the work;

f) **iterative questioning.** In addition to the “preventative” strategies outlined above, specific ploys may be incorporated to uncover deliberate lies. These might include the use of probes to elicit detailed data and iterative questioning, in which the researcher returns to matters previously raised by an informant and extracts related data through rephrased questions. In both cases, where contradictions emerge, falsehoods can be detected and the researcher may decide to discard the suspect data. An alternative approach and one that provides greater transparency lies in drawing attention, within the final research report, to the discrepancies and offering possible explanations;

g) **negative case analysis**, as recommended by commentators such as Lincoln and Guba [25], Miles and Huberman [26] and Silverman [27]. One form of negative case analysis may see the researcher refining a hypothesis until it addresses all cases within the data. If the study includes the production of typologies, on completing the initial categories the investigator may revisit the data in order to confirm that these constructs do indeed account for all instances of the phenomenon involved, even if some of the types embrace only one instance;

h) **frequent debriefing sessions** between the researcher and his or her superiors, such as a project director or steering group. Through discussion, the vision of the investigator may be widened as others bring to bear their experiences and perceptions. Such collaborative sessions can be used by the researcher to discuss alternative approaches, and others who are responsible for the work in a more supervisory capacity may draw attention to flaws in the proposed course of action. The meetings also provide a sounding board for the investigator to test his or her developing ideas and interpretations, and probing from others may help the researcher to recognise his or her own biases and preferences;

i) **peer scrutiny of the research project.** Opportunities for scrutiny of the project by colleagues, peers and academics should be welcomed, as should feedback offered to the researcher at any presentations (e.g. at conferences) that are made over the duration of the project. The fresh perspective that such individuals may be able to bring may allow them to challenge assumptions made by the investigator, whose closeness to the project frequently inhibits his or her ability to view it with real detachment. Questions and observations may well enable the researcher to refine his or her methods, develop a greater explanation of the research design and strengthen his or her arguments in the light of the comments made;
j) the researcher’s “reflective commentary”. In addition to the outside scrutiny discussed above, the investigator should seek to evaluate the project, again as it develops. This may be done through a reflective commentary, part of which may be devoted to the effectiveness of the techniques that have been employed. The reflective commentary may also be used to record the researcher’s initial impressions of each data collection session, patterns appearing to emerge in the data collected and theories generated. The commentary can play a key role in what Guba and Lincoln term “progressive subjectivity”, or the monitoring of the researcher’s own developing constructions, which the writers consider critical in establishing credibility [28]. Ultimately, the section of the commentary dealing with emerging patterns and theories should inform that part of the research report that addresses the project’s results, and any discussion in the report of the effectiveness of the study may be based on the investigator’s methods analysis within the reflective commentary;

k) background, qualifications and experience of the investigator. According to Patton, the credibility of the researcher is especially important in qualitative research as it is the person who is the major instrument of data collection and analysis [29]. Alkin, Daillak and White go so far as to suggest that a scrutineer’s trust in the researcher is of equal importance to the adequacy of the procedures themselves [30].

The nature of the biographical information that should be supplied in the research report is a matter of debate. Maykut and Morehouse recommend including any personal and professional information relevant to the phenomenon under study [31], and Patton adds that arrangements by which the investigator is funded should also be addressed [32]. Any approvals given to the project by those providing access to the organisation and individual participants should also be made explicit;

l) member checks, which Guba and Lincoln consider the single most important provision that can be made to bolster a study’s credibility [33]. Checks relating to the accuracy of the data may take place “on the spot” in the course, and at the end, of the data collection dialogues. Informants may also be asked to read any transcripts of dialogues in which they have participated. Here the emphasis should be on whether the informants consider that their words match what they actually intended, since, if a tape recorder has been used, the articulations themselves should at least have been accurately captured.

Another element of member checking should involve verification of the investigator’s emerging theories and inferences as these were formed during the dialogues. This strategy has been employed by Pitts [34] and is recommended by Brewer and Hunter [35] and Miles and Huberman [36]. Where appropriate, participants may be asked if they can offer reasons for particular patterns observed by the researcher. The importance of developing such a formative understanding is recognised by Van Maanen, who writes that
“analysis and verification... is something one brings forth with them from the field, not something which can be attended to later, after the data are collected. When making sense of field data, one cannot simply accumulate information without regard to what each bit of information represents in terms of its possible contextual meanings” [37];

m) *thick description of the phenomenon under scrutiny.* Detailed description in this area can be an important provision for promoting credibility as it helps to convey the actual situations that have been investigated and, to an extent, the contexts that surround them. Without this insight, it is difficult for the reader of the final account to determine the extent to which the overall findings “ring true”. Moreover, if the researcher employs a reporting system in which he or she defines a series of types within a typology and illustrates these types using real qualitative episodes, the inclusion of the latter enables the reader to assess how far the defined types truly embrace the actual situations;

n) *examination of previous research findings* to assess the degree to which the project’s results are congruent with those of past studies. Silverman considers that the ability of the researcher to relate his or her findings to an existing body of knowledge is a key criterion for evaluating works of qualitative inquiry [38]. In this respect, reports of previous studies staged in the same or a similar organisation and addressing comparable issues may be invaluable sources.

3. Transferability

Merriam writes that external validity “is concerned with the extent to which the findings of one study can be applied to other situations” [39]. In positivist work, the concern often lies in demonstrating that the results of the work at hand can be applied to a wider population. Since the findings of a qualitative project are specific to a small number of particular environments and individuals, it is impossible to demonstrate that the findings and conclusions are applicable to other situations and populations. Erlandson et al. note that many naturalistic inquirers believe that, in practice, even conventional generalisability is never possible as all observations are defined by the specific contexts in which they occur [40]. A contrasting view is offered by Stake [41] and Denscombe [42], who suggest that, although each case may be unique, it is also an example within a broader group and, as a result, the prospect of transferability should not be immediately rejected. Nevertheless, such an approach can be pursued only with caution since, as Gomm, Hammersley and Foster recognise, it appears to belittle the importance of the contextual factors which impinge on the case [43].

Bassey proposes that, if practitioners believe their situations to be similar to that described in the study, they may relate the findings to their own positions [44]. Lincoln and Guba [45] and Firestone [46] are among those who present a similar argument, and suggest that it is the responsibility of the investigator to ensure that sufficient contextual information about the fieldwork sites is provided to enable the
reader to make such a transfer. They maintain that, since the researcher knows only the “sending context”, he or she cannot make transferability inferences. In recent years such a stance has found favour with many qualitative researchers. After perusing the description within the research report of the context in which the work was undertaken, readers must determine how far they can be confident in transferring to other situations the results and conclusions presented. It is also important that sufficient thick description of the phenomenon under investigation is provided to allow readers to have a proper understanding of it, thereby enabling them to compare the instances of the phenomenon described in the research report with those that they have seen emerge in their situations.

Authors disagree on the nature and extent of background information that should be offered but few would dispute the need for “a full description of all the contextual factors impinging on the inquiry”, as recommended by Guba [47]. Nevertheless, the situation is complicated by the possibility, noted by Firestone, that factors considered by the researcher to be unimportant, and consequently unaddressed in the research report, may be critical in the eyes of a reader [48]. Many investigators stop short of the course of action advocated by Denscombe that the researcher should demonstrate how, in terms of the contextual data, the case study location(s) compare(s) with other environments [49]. This reluctance is based on the fact that the process would demand a considerable knowledge of the “receiving contexts” of other organisations, and the researcher is in no position to comment on what Merriam calls the “typicality” of the environment(s) in which the fieldwork took place [50].

The work of Cole and Gardner [51], Marchionini and Teague [52] and Pitts [53] highlights the importance of the researcher’s conveying to the reader the boundaries of the study. This additional information must be considered before any attempts at transference are made. Thus information on the following issues should be given at the outset:

a) the number of organisations taking part in the study and where they are based;

b) any restrictions in the type of people who contributed data;

c) the number of participants involved in the fieldwork;

d) the data collection methods that were employed;

e) the number and length of the data collection sessions;

f) the time period over which the data was collected.

It is easy for researchers to develop a preoccupation with transferability. Ultimately, the results of a qualitative study must be understood within the context of the particular characteristics of the organisation or organisations and, perhaps, geographical area in which the fieldwork was carried out. In order to assess the extent to which findings may be true of people in other settings, similar projects employing the same methods but conducted in different environments could well be of great value. As Kuhlthau [54] and Gomm, Hammersley and Foster [55] recognise, however, it is rare for such complementary work to be undertaken. Nevertheless, the accumulation of findings from studies staged in different settings might enable a more inclusive,
overall picture to be gained. A similar point is made by Gross, in relation to her work on imposed queries in school libraries. She writes of the “multiple environments” in which the phenomenon of her interest takes place and believes her study to provide a “baseline understanding” with which the results of subsequent work should be compared [56]. As Borgman [57] and Pitts [58] have acknowledged, understanding of a phenomenon is gained gradually, through several studies, rather than one major project conducted in isolation. Even when different investigations offer results that are not entirely consistent with one another, this does not, of course, necessarily imply that one or more is untrustworthy. It may be that they simply reflect multiple realities, and, if an appreciation can be gained of the reasons behind the variations, this understanding may prove as useful to the reader as the results actually reported. Such an attitude is consistent with what Dervin considers should be key principles within information-seeking research, namely: “To posit... every contradiction, every inconsistency, every diversity not as an error or extraneous but as fodder for contextual analysis. To ask and re-ask what accounts for this difference or this similarity and to anchor possible answers in time-space conceptualizings” [59]. It should thus be questioned whether the notion of producing truly transferable results from a single study is a realistic aim or whether it disregards the importance of context which forms such a key factor in qualitative research.

4. Dependability

In addressing the issue of reliability, the positivist employs techniques to show that, if the work were repeated, in the same context, with the same methods and with the same participants, similar results would be obtained. However, as Fidel [60] and Marshall and Rossman [61] note, the changing nature of the phenomena scrutinised by qualitative researchers renders such provisions problematic in their work. Florio-Ruane highlights how the investigator’s observations are tied to the situation of the study, arguing that the “published descriptions are static and frozen in the ‘ethnographic present’ ” [62]. Lincoln and Guba stress the close ties between credibility and dependability, arguing that, in practice, a demonstration of the former goes some distance in ensuring the latter [63]. This may be achieved through the use of “overlapping methods”, such as the focus group and individual interview.

In order to address the dependability issue more directly, the processes within the study should be reported in detail, thereby enabling a future researcher to repeat the work, if not necessarily to gain the same results. Thus, the research design may be viewed as a “prototype model”. Such in-depth coverage also allows the reader to assess the extent to which proper research practices have been followed. So as to enable readers of the research report to develop a thorough understanding of the methods and their effectiveness, the text should include sections devoted to

a) *the research design and its implementation*, describing what was planned and executed on a strategic level;
b) the operational detail of data gathering, addressing the minutiae of what was done in the field;
c) reflective appraisal of the project, evaluating the effectiveness of the process of inquiry undertaken.

5. Confirmability

Patton associates objectivity in science with the use of instruments that are not dependent on human skill and perception. He recognises, however, the difficulty of ensuring real objectivity, since, as even tests and questionnaires are designed by humans, the intrusion of the researcher’s biases is inevitable [64]. The concept of confirmability is the qualitative investigator’s comparable concern to objectivity. Here steps must be taken to help ensure as far as possible that the work’s findings are the result of the experiences and ideas of the informants, rather than the characteristics and preferences of the researcher. The role of triangulation in promoting such confirmability must again be emphasised, in this context to reduce the effect of investigator bias. Miles and Huberman consider that a key criterion for confirmability is the extent to which the researcher admits his or her own predispositions [65]. To this end, beliefs underpinning decisions made and methods adopted should be acknowledged within the research report, the reasons for favouring one approach when others could have been taken explained and weaknesses in the techniques actually employed admitted. In terms of results, preliminary theories that ultimately were not borne out by the data should also be discussed. Much of the content in relation to these areas may be derived from the ongoing “reflective commentary”.

Once more, detailed methodological description enables the reader to determine how far the data and constructs emerging from it may be accepted. Critical to this process is the “audit trail”, which allows any observer to trace the course of the research step-by-step via the decisions made and procedures described. The “audit trail” may be represented diagrammatically. Two such diagrams may be constructed. One may take a data-oriented approach, showing how the data eventually leading to the formation of recommendations was gathered and processed during the course of the study. This is what is typically understood by the term, “audit trail”. In addition, however, the manner in which the concepts inherent in the research question gave rise to the work to follow may be tracked. This more theoretical “audit trail”, which should be understood in terms of the whole of the duration of the project, may be depicted in a second diagram.

6. Summary and conclusion

Over the last twenty years, much has been achieved by advocates of qualitative inquiry in demonstrating the rigour and trustworthiness of their favoured form of
Provisions that may be Made by a Qualitative Researcher Wishing to Address Guba’s Four Criteria for Trustworthiness

<table>
<thead>
<tr>
<th>Quality criterion</th>
<th>Possible provision made by researcher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credibility</td>
<td>Adoption of appropriate, well recognised research methods</td>
</tr>
<tr>
<td></td>
<td>Development of early familiarity with culture of participating organisations</td>
</tr>
<tr>
<td></td>
<td>Random sampling of individuals serving as informants</td>
</tr>
<tr>
<td></td>
<td>Triangulation via use of different methods, different types of informants and different sites</td>
</tr>
<tr>
<td></td>
<td>Tactics to help ensure honesty in informants</td>
</tr>
<tr>
<td></td>
<td>Iterative questioning in data collection dialogues</td>
</tr>
<tr>
<td></td>
<td>Negative case analysis</td>
</tr>
<tr>
<td></td>
<td>Debriefing sessions between researcher and superiors</td>
</tr>
<tr>
<td></td>
<td>Peer scrutiny of project</td>
</tr>
<tr>
<td></td>
<td>Use of “reflective commentary”</td>
</tr>
<tr>
<td></td>
<td>Description of background, qualifications and experience of the researcher</td>
</tr>
<tr>
<td></td>
<td>Member checks of data collected and interpretations/theories formed</td>
</tr>
<tr>
<td></td>
<td>Thick description of phenomenon under scrutiny</td>
</tr>
<tr>
<td></td>
<td>Examination of previous research to frame findings</td>
</tr>
<tr>
<td>Transferability</td>
<td>Provision of background data to establish context of study and detailed description of phenomenon in question to allow comparisons to be made</td>
</tr>
<tr>
<td>Dependability</td>
<td>Employment of “overlapping methods”</td>
</tr>
<tr>
<td></td>
<td>In-depth methodological description to allow study to be repeated</td>
</tr>
<tr>
<td>Confirmability</td>
<td>Triangulation to reduce effect of investigator bias</td>
</tr>
<tr>
<td></td>
<td>Admission of researcher’s beliefs and assumptions</td>
</tr>
<tr>
<td></td>
<td>Recognition of shortcomings in study’s methods and their potential effects</td>
</tr>
<tr>
<td></td>
<td>In-depth methodological description to allow integrity of research results to be scrutinised</td>
</tr>
<tr>
<td></td>
<td>Use of diagrams to demonstrate “audit trail”</td>
</tr>
</tbody>
</table>

research. Nevertheless, criticisms of work of this kind continue to be made by positivists. This paper has addressed four criteria that may be addressed by qualitative researchers wishing to present a convincing case that their work is academically sound. A range of strategies that may be adopted by investigators in response to these issues has been highlighted. These are summarised in the chart below:

The challenge for those involved in teaching courses in research methods lies in ensuring that those contemplating undertaking qualitative research are not only aware of the criticisms typically made by its detractors but they are also cognisant of the provisions which can be made to address matters such as credibility, transferability, dependability and confirmability. Prospective researchers can then assess the extent to which they are able to apply these generic strategies to their particular investigation.

References


[33] J.M. Pitts, op. cit.

[34] J. Brewer and A. Hunter, op. cit.
[58] J.M. Pitts, op. cit.
[64] M.Q. Patton, op. cit.