Rigor in Qualitative Research: The Assessment of Trustworthiness

Laura Krefting

Key Words: research • research design • social sciences

Despite a growing interest in qualitative research in occupational therapy, little attention has been placed on establishing its rigor. This article presents one model that can be used for the assessment of trustworthiness or merit of qualitative inquiry. Cuba's (1981) model describes four general criteria for evaluation of research and then defines each from both a quantitative and a qualitative perspective. Several strategies for the achievement of rigor in qualitative research useful for both researchers and consumers of research are described.

The worth of any research endeavor, regardless of the approach taken, is evaluated by peers, grant reviewers, and readers. Most quantitative researchers recognize and document the worth of a project by assessing the reliability and validity of the work (Payton, 1979). This same attention to the merit of a research project, however, is much less common in qualitative research.

Schmid (1981) described qualitative research as the study of the empirical world from the viewpoint of the person under study. She identified two underlying principles. The first is that behavior is influenced by the physical, sociocultural, and psychological environment—this is the basis for naturalistic inquiry. The second assumption is that behavior goes beyond what is observed by the investigator. Subjective meanings and perceptions of the subject are critical in qualitative research, and it is the researcher's responsibility to access these.

Kirk and Miller (1986) suggested a working definition of qualitative research that reflects these two principles. They defined it as "a particular tradition in social science that fundamentally depends on watching people in their own territory and interacting with them in their own language, on their own terms" (p. 9). Qualitative research is pluralistic, consisting of a variety of approaches, including phenomenology, semiotics, ethnography, life history, and historical research. Detailed definitions of the epistemological and philosophical bases of qualitative research reviews can be found in Duffy (1985), Guba (1981), and Kielhofner (1982).

The purpose of this paper is to describe one conceptual model that may be used to evaluate the trustworthiness of qualitative research. Four concepts basic to the model are described, followed by strategies that researchers can use to enhance the worth of their qualitative studies.

Two issues need to be acknowledged in a discussion of the assessment of qualitative research. The first is that models used to evaluate quantitative research are seldom relevant to qualitative research. The second is that not all qualitative research can be assessed with the same strategies. These two issues will be addressed briefly before the model is presented.

Criteria for Assessment of Qualitative Versus Quantitative Research

Too frequently, qualitative research is evaluated against criteria appropriate to quantitative research and is found to be lacking. Qualitative researchers contend that because the nature and purpose of the quantitative and qualitative traditions are different, it is erroneous to apply the same criteria of worthiness or merit. Agar (1986) suggested that terms like reliability and validity are relative to the quantitative view and do not fit the details of qualitative research. For example, the notion of external...
validity, which is concerned with the ability to generalize from the research sample to the population (Payton, 1979), is one of the key criteria of good quantitative research. In some qualitative approaches, however, the major purpose is to generate hypotheses for further investigation rather than to test them (Sandelowski, 1986). In such situations, external validity is not relevant.

Agar (1986) suggested that a different language is needed to fit the qualitative view, one that would replace reliability and validity with such terms as acceptability, accuracy of representation, and authorizy of the writer. Similarly, Leininger (1985) claimed that the issue is not whether the data are reliable or valid but how the terms reliability and validity are defined. She recast the term validity in a qualitative sense to mean gaining knowledge and understanding of the nature (i.e., the meaning, attributes, and characteristics) of the phenomenon under study. She contrasted this to the common usage of validity in a quantitative sense, in which it refers to the degree to which an instrument measures what it is designed to measure.

Just as there is a need to look at the accuracy and trustworthiness of various kinds of quantitative data in different ways, there is also a need to look at qualitative methods for the different ways in which to ensure the quality of the findings. It is important not to fall into the trap of assuming that all qualitative studies should be evaluated with the same criteria. As Sandelowski (1986) noted, the term qualitative research is imprecise and refers to many dissimilar research methods. These approaches have different purposes and methods and therefore different ways of determining whether they are trustworthy. For example, the phenomenological approach asks what it is like to have a certain experience. The goal is to describe accurately the experience of the phenomenon under study, not to generalize to theories or models (Field & Morse, 1985). Because the goal of ethnographic research, however, is to describe social complexities, it often involves development of theoretical constructs. Although some principles are basic to all qualitative research, the incorrect application of the qualitative criteria of trustworthiness to studies is as problematic as the application of inappropriate quantitative criteria.

The issues of the inappropriateness of quantitative criteria in the assessment of qualitative research and the plurality of qualitative research are important to the understanding of any model of trustworthiness to studies as is problematic as the application of inappropriate quantitative criteria.

Guba’s Model of Trustworthiness of Qualitative Research

Researchers need alternative models appropriate to qualitative designs that ensure rigor without sacrificing the relevance of the qualitative research. Guba (1981) proposed such a model for assessing the trustworthiness of qualitative data. Although there are other models (e.g., Kirk & Miller, 1986; Leininger, 1985), this presentation will be based on Guba’s model because it is comparatively well developed conceptually and has been used by qualitative researchers, particularly nurses and educators, for a number of years.

What follows is my summary and interpretation of Guba’s (1981) model following several qualitative research projects. I urge occupational therapists who intend to make use of the model in designing research projects or in the critical assessment of the projects of others to refer to the original sources cited in the references.

Guba’s (1981) model is based on the identification of four aspects of trustworthiness that are relevant to both quantitative and qualitative studies: (a) truth value, (b) applicability, (c) consistency, and (d) neutrality. Based on the philosophical differences between qualitative and quantitative approaches, the model defines different strategies of assessing these criteria in each type of research. These strategies are important to researchers in designing ways of increasing the rigor of their qualitative studies and also for readers to use as a means of assessing the value of the findings of qualitative research.

Truth Value

Truth value asks whether the researcher has established confidence in the truth of the findings for the subjects or informants and the context in which the study was undertaken (Lincoln & Guba, 1985). It establishes how confident the researcher is with the truth of the findings based on the research design, informants, and context. In quantitative studies, truth is often assessed by how well threats to the internal validity of the study have been managed as well as the validity of the instruments as a measure of the phenomenon under study (Sandelowski, 1986). Internal validity is supported when changes in the dependent variable are accounted for by changes in the independent variable, that is, when the design minimizes the effects of competing confounding variables by control or randomization. (For a detailed discussion of the threats to internal validity, see Campbell and Stanley, 1966).

In qualitative research, truth value is usually obtained from the discovery of human experiences as they are lived and perceived by informants. Truth value is subject-oriented, not defined a priori by the researcher (Sandelowski, 1986). Lincoln and Guba (1985) termed this credibility. They argued that internal validity is based on the assumption that there is a single tangible reality to be measured. If this assumption is replaced by the idea of multiple realities, the researcher’s job becomes one of representing these multiple realities revealed by informants as adequately as possible. Researchers, then, need to focus on testing their findings against various groups.
from which the data were drawn or persons who are familiar with the phenomenon being studied. Sandelowski suggested that a qualitative study is credible when it presents such accurate descriptions or interpretation of human experience that people who also share that experience would immediately recognize the descriptions. Truth value is perhaps the most important criterion for the assessment of qualitative research. A number of methodological strategies are required to ensure strong credibility.

**Applicability**

Applicability refers to the degree to which the findings can be applied to other contexts and settings or with other groups; it is the ability to generalize from the findings to larger populations. In the quantitative perspective, applicability refers to how well the threats to external validity have been managed (Sandelowski, 1986). Payton (1979) defined external validity as the ability to generalize from the study sample to the larger population and noted the importance of sampling technique in its establishment.

Two perspectives to applicability are appropriate for qualitative research. The first perspective suggests that the ability to generalize is not relevant in many qualitative research projects. A strength of the qualitative method is that it is conducted in naturalistic settings with few controlling variables. Each situation is defined as unique and thus is less amenable to generalization. Consequently, as Sandelowski (1986) explained, generalization is somewhat of an illusion because every research situation is made up of a particular researcher in a particular interaction with particular informants. Applicability, then, is not seen as relevant to qualitative research because its purpose is to describe a particular phenomenon or experience, not to generalize to others.

Guba (1981) presented the second perspective on applicability in qualitative research by referring to **fittingness**, or transferability, as the criterion against which applicability of qualitative data is assessed. Research meets this criterion when the findings fit into contexts outside the study situation that are determined by the degree of similarity or goodness of fit between the two contexts. Lincoln and Guba (1985) noted that transferability is more the responsibility of the person wanting to transfer the findings to another situation or population than that of the researcher of the original study. They argued that as long as the original researcher presents sufficient descriptive data to allow comparison, he or she has addressed the problem of applicability.

**Consistency**

The third criterion of trustworthiness considers the consistency of the data, that is, whether the findings would be consistent if the inquiry were replicated with the same subjects or in a similar context. In quantitative research, **reliability** is the criterion concerned with the stability, consistency, and equivalence in the study (Sandelowski, 1986). It is the extent to which repeated administration of a measure will provide the same data or the extent to which a measure administered once, but by different people, produces equivalent results. Inherent in the goal of reliability is the value of repeatability, that replication of the testing procedures does not alter the findings. The restricted methods of observation and tight designs common in quantitative research are intended to pass this replication test. The quantitative perspective on consistency is also based on the assumption of a single reality, that there is something out there to be studied that is unchanging and can be used as a benchmark (Lincoln & Guba, 1985). If one assumes there are multiple realities, the notion of reliability is no longer as relevant.

Unlike the relatively controlled experimental environment, the qualitative field setting may be complicated by extraneous and unexpected variables. As Duffy (1985) noted, the structure of the experimental design is the antithesis of the unstructured and often spontaneous strategies of qualitative research. The key to qualitative work is to learn from the informants rather than control for them. Moreover, instruments that are assessed for consistency in qualitative research are the researcher and the informants, both of whom vary greatly within the research project. Qualitative research emphasizes the uniqueness of the human situation, so that variation in experience rather than identical repetition is sought (Field & Morse, 1985).

Thus, variability is expected in qualitative research, and consistency is defined in terms of dependability. Guba’s (1981) concept of dependability implies trackable variability, that is, variability that can be ascribed to identifiable sources. Explainable sources of variability might include increasing insight on the part of the researcher, informant fatigue, or changes in the informant’s life situation. Another source of variability stems from the fact that qualitative research looks at the range of experience rather than the average experience, so that atypical or non-normative situations are important to include in the findings. In quantitative terms, the outlying data need to be identified to describe the boundaries of the experience or phenomenon. Although the person might not be completely representative of a group, his or her experience is considered important.

**Neutrality**

The fourth criterion of trustworthiness is neutrality, the freedom from bias in the research procedures and results (Sandelowski, 1986). Neutrality refers to the degree to which the findings are a function solely of the informants and conditions of the research and not of other biases, motivations, and perspectives (Guba, 1981). In quantita-
tive research, objectivity is the criterion of neutrality and is achieved through rigor of methodology through which reliability and validity are established. Objectivity also refers to the proper distance between researcher and subjects that minimizes bias and is achieved through such procedures as instrumentation and randomization. Thus, the objective researcher is seen as scientifically distant, that is, as someone who is not influenced by, and does not influence, the study.

Qualitative researchers, on the other hand, try to increase the worth of the findings by decreasing the distance between the researcher and the informants, for example, by prolonged contact with informants or lengthy periods of observation. Lincoln and Guba (1985) shifted the emphasis of neutrality in qualitative research from the researcher to the data, so that rather than looking at the neutrality of the investigator, the neutrality of the data was considered. They suggested that confirmability be the criterion of neutrality. This is achieved when truth value and applicability are established.

Summary of Guba's Model

Guba's (1981) model identified truth value, applicability, consistency, and neutrality as four criteria applicable to the assessment of research of any type. Guba argued that these criteria must be defined differently for qualitative and quantitative research based on the philosophical and conceptual divergence of the two approaches. (For a summary of the criteria, their common quantitative definitions, and Guba's qualitative definitions, see Table 1.)

Strategies to Increase the Trustworthiness of Qualitative Work

Specific strategies can be used throughout the research process to increase the worth of qualitative projects. Some strategies need to be addressed in the study design stage, while others are addressed during data collection and after data are interpreted. A small number of these strategies will be discussed in detail, such as reflexivity and triangulation, because they are critical to the quality of the research, while other, more straightforward strategies will be described briefly. The strategies are described under one of the four qualitative criteria for trustworthiness. Although some strategies are useful for establishing more than one criterion (e.g., triangulation and reflexivity), the strategies are defined under the criterion to which they are most frequently applied. Many of these strategies are outlined in Guba's (1981) and Lincoln and Guba's (1985) studies. I learned about other strategies, however, from working with experienced researchers, and a few were developed to strengthen my own work. (See Table 2 for a summary of the strategies according to the relevant criteria.) In reviewing these strategies, it is important to remember that although a number of techniques are available, not all are appropriate to every qualitative study.

Credibility Strategies

Leininger (1985) noted the importance of identifying and documenting recurrent features such as patterns, themes, and values in qualitative research. The emphasis on recurrence suggests the need to spend sufficient time with informants to identify reappearing patterns. Credibility requires adequate submersion in the research setting to enable recurrent patterns to be identified and verified. Thus, an important strategy is to spend an extended period of time with informants (Lincoln and Guba [1985] termed this prolonged engagement), which allows the researcher to check perspectives and allows the informants to become accustomed to the researcher. Kielhofner (1982) supported the importance of intense participation, suggesting that it enhances research findings through intimate familiarity and discovery of hidden fact. This extended time period is important because as rapport increases, informants may volunteer different

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Comparison of Criteria by Research Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion</td>
<td>Qualitative Approach</td>
</tr>
<tr>
<td>Truth value</td>
<td>Credibility</td>
</tr>
<tr>
<td>Applicability</td>
<td>Transferability</td>
</tr>
<tr>
<td>Consistency</td>
<td>Dependability</td>
</tr>
<tr>
<td>Neutrality</td>
<td>Confirmability</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Summary of Strategies With Which to Establish Trustworthiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy</td>
<td>Credibility</td>
</tr>
<tr>
<td>Prolonged and varied field experience</td>
<td></td>
</tr>
<tr>
<td>Time sampling</td>
<td></td>
</tr>
<tr>
<td>Reflexivity (field journal)</td>
<td></td>
</tr>
<tr>
<td>Triangulation</td>
<td></td>
</tr>
<tr>
<td>Member checking</td>
<td></td>
</tr>
<tr>
<td>Peer examination</td>
<td></td>
</tr>
<tr>
<td>Establishing authority of researcher</td>
<td></td>
</tr>
<tr>
<td>Structural coherence</td>
<td></td>
</tr>
<tr>
<td>Referential adequacy</td>
<td></td>
</tr>
<tr>
<td>Transferability</td>
<td>Nominated sample</td>
</tr>
<tr>
<td>Comparison of sample to demographic data</td>
<td></td>
</tr>
<tr>
<td>Time sample</td>
<td></td>
</tr>
<tr>
<td>Dense description</td>
<td></td>
</tr>
<tr>
<td>Dependability</td>
<td>Dependability audit</td>
</tr>
<tr>
<td>Dense description of research methods</td>
<td></td>
</tr>
<tr>
<td>Stepwise replication</td>
<td></td>
</tr>
<tr>
<td>Triangulation</td>
<td></td>
</tr>
<tr>
<td>Peer examination</td>
<td></td>
</tr>
<tr>
<td>Code-recode procedure</td>
<td></td>
</tr>
<tr>
<td>Confirmability</td>
<td>Confirmability audit</td>
</tr>
<tr>
<td>Triangulation</td>
<td></td>
</tr>
<tr>
<td>Reflexivity</td>
<td></td>
</tr>
</tbody>
</table>
and often more sensitive information than they do at the beginning of a research project.

There are no rules regulating the length of time one should be involved in data collection. It depends on the design and the particular purpose of the study. With the use of my experience as a guide, my disability ethnography (Krefting, 1989) consisted of more than eighty 90-min interviews and 20 to 30 periods of participant observations with 22 persons with brain injury and their families. A life history, which is in progress, consists of more than 20 hr of interviews plus 50 pages of autobiographical writing.

A study’s credibility is threatened by errors in which research subjects respond with what they think is the preferred social response—that is, data are based on social desirability rather than on personal experience (Kirk & Miller, 1986). Use of prolonged engagement can assist in detecting response sets where informants consistently either agree or disagree with the questions. Use of numerous interviews and observation periods allows the researcher to identify the occurrence of this problem. The use of hypothetical cases or a reframing of the questions may help elicit more personal responses.

Related is the issue of persistent observation of a phenomenon under various natural situations. The time-sampling strategy makes use of a flowchart to systematize informant contacts and observation to determine if the researcher is sampling all possible situations, including different social settings: times of day, week, and season; and interactions among different social groupings (Knafl & Breitmayer, 1989). For example, in studying the lives of children with disabilities, the researcher would want to observe the children interacting with peers, family, and teachers in a number of contexts at different times of the day and on weekends and weekdays. This strategy emphasizes the importance of the environment in which the data are collected and establishes credibility.

Paradoxically, a major threat to the truth value of a qualitative study lies in the closeness of the relationship between the investigator and the informants that can develop during the prolonged contact required to establish credibility. The researcher can become so enmeshed with the informants that he or she may have difficulty separating his or her own experience from that of the informants (Marcus & Fischer, 1986). Although close researcher-informant relationships are critical to the research enterprise, it is possible to lose the ability to interpret the findings. To help ensure that this extreme overinvolvement does not occur, a strategy called reflexive analysis, or reflexivity, is useful (Good, Herrera, Good, & Cooper, 1985).

Reflexivity refers to assessment of the influence of the investigator’s own background, perceptions, and interests on the qualitative research process (Ruby, 1980). It includes the effect of the researcher’s personal history on qualitative research. In the past, many qualitative researchers have claimed neutrality and even invisibility in their fieldwork, much as the objective scientist does in the quantitative approach to research. The focus on reflexivity is a recent trend in cultural anthropology, as evidenced in the work of Crapanzano (1980) in Morocco and Shostak (1981) with the !Kung tribe. Agar (1986) noted that the researcher’s background dictates the framework from which he or she will organize, study, and analyze the findings. This background is made up of all of the resources available to make sense out of the experience and is often reflected in multiple roles the researcher plays while engaged in the research. For example, in my ethnographic study of persons with traumatic brain injury (Krefting, 1989), I maintained a double role as a clinical anthropology student and as an experienced health care professional. As an anthropologist, I wanted to do pure ethnography, that is, see community life after brain injury only from the viewpoint of the informants, without the influence of my rehabilitation training. But as a health care provider, I was trained to help, which motivated me to look for specific and practical implications in what I saw and heard. I spent most of the study juggling the influences of the two roles. It was important for me to be aware of and to reflect on the influence of these two roles on the study.

Aamodt (1982) noted that the qualitative approach is reflexive in that the researcher is part of the research, not separate from it. Research situations are dynamic, and the researcher is a participant, not merely an observer. The investigator, then, must analyze himself or herself in the context of the research. As an anthropologist, I maintained a double role as a clinical anthropology student and as an experienced health care professional. As an anthropologist, I wanted to do pure ethnography, that is, see community life after brain injury only from the viewpoint of the informants, without the influence of my rehabilitation training. But as a health care provider, I was trained to help, which motivated me to look for specific and practical implications in what I saw and heard. I spent most of the study juggling the influences of the two roles. It was important for me to be aware of and to reflect on the influence of these two roles on the study.

One of the ways that researchers can describe and interpret their own behavior and experiences within the research context is to make use of a field journal. This journal is kept throughout the research process and includes three types of information (Lincoln & Guba, 1985). The daily schedule and logistics of the study and methods log (in which decisions about methods and their rationale are described) are two components of the field journal that are important for auditability, a strategy that will be discussed in a later section. The third, and most important, type of information in the field journal is analogous to that found in a personal diary and reflects the researcher’s thoughts, feelings, ideas, and hypotheses generated by contact with informants. It also contains questions, problems, and frustrations concerning the overall research process. In writing these personal thoughts and feelings about the research process, the researcher may become aware of biases and preconceived assumptions. Once aware of these biases, the researcher may alter the way that he or she collects the data.
or approaches the analysis to enhance the credibility of the research.

Triangulation is a powerful strategy for enhancing the quality of the research, particularly credibility. It is based on the idea of convergence of multiple perspectives for mutual confirmation of data to ensure that all aspects of a phenomenon have been investigated (Knafl & Breitmayer, 1989). The triangulated data sources are assessed against one another to cross-check data and interpretation. This strategy of providing a number of different slices of data also minimizes distortion from a single data source or from a biased researcher, as may be the case in data based on a single application of one measure, a single client interview for example (Field & Morse, 1985). Researchers and readers need to consider how the triangulation either contributed to confirmation of certain aspects of the study or to the completeness with which the phenomenon of interest was addressed.

A number of types of triangulation exist, four of which were identified by Knafl and Breitmayer (1989). The most common is triangulation of data methods, in which data collected by various means are compared (e.g., data from structured interviews, participant observation, life histories). A second type, triangulation of data sources, maximizes the range of data that might contribute to complete understanding of the concept. It is based on the importance of variety in time, space, and person in observation and interviewing. Examples of triangulated sources include different seasons or days, different settings, and different groupings of people. In a hospital setting, one might observe different shifts on different wards and focus on family, patients, and professionals, either alone or in small groups. Theoretical triangulation means that ideas from diverse or competing theories can be tested. For example, triangulation of theory may occur in considering a number of concepts from anthropology, rehabilitation, sociology, and psychology in the conceptual interpretation of the experience of a disability. Triangulation of investigators occurs in a study in which a research team, rather than a single researcher, is used. Team members often have a diversity of approaches, for example, a team investigating the costs of home care for patients with arthritis may be made up of a physician, a therapist, an anthropologist, and an economist.

Central to the credibility of qualitative research is the ability of informants to recognize their experiences in the research findings. Member checking is a technique that consists of continually testing with informants the researcher’s data, analytic categories, interpretations, and conclusions (Lincoln & Guba, 1985). This strategy of revealing research materials to the informants ensures that the researcher has accurately translated the informants’ viewpoints into data. Assessment to see if the data make sense through member checking decreases the chances of misrepresentation.

Member checking can be done through an interweaving of the informant contact hours so that information from one interview is checked with another informant before a subsequent interview with the first. In addition, summaries of taped interviews can be played to informants for their responses, or more formalized work sessions can be held, in which a number of informants are gathered to react to a draft of the analytical codes or report of results, for example. To test the overall interpretation, near the conclusion of the study, one must do a terminal or final member check with key informants to ensure that the final presentation of the data reflects the experience accurately (Lincoln & Guba, 1985). Member checks are more difficult for informants to carry out at the latter stages of the research process, when higher conceptual analysis is necessary, than in the data gathering phase, in which descriptive data are reviewed by informants. The selection criteria for informants for a terminal member check, therefore, are critical.

Despite the usefulness of member checking in enhancing credibility, one must consider the ethical aspect of this strategy. Researchers must be selective about which informants are involved in member checking. Often, informants are not conscious of the information discovered by the researcher and may become troubled if made aware of it. For example, during participant observation, the researcher may have noted that the spouses of persons with disabilities treat them like children. The person who is involved in member checking these data should be carefully selected to ensure that he or she would not be harmed by reading the data. The researcher should not provide insight that might, in fact, be harmful to the well-being of informants. Another difficulty with member checks is that informants may have a tendency to internalize the information they have read, which could affect their subsequent responses. To minimize this, it is best not to reinterview or observe an informant on an aspect of the project for which he or she has conducted a member check.

Peer examination is based on the same principle as member checks but involves the researcher’s discussing the research process and findings with impartial colleagues who have experience with qualitative methods. Insights are discussed and problems presented as a form of debriefing. Lincoln and Guba (1985) suggested that this is one way of keeping the researcher honest, and the searching questions may contribute to deeper reflexive analysis by the researcher. Colleagues can also increase credibility by checking categories developed out of data and by looking for disconfirming or negative cases. Peer examination also presents an opportunity for the researcher to present working hypotheses for reaction and to discuss the evolving design of the study (Lincoln &
The availability of the informants' verbatim accounts (tape recordings or transcripts of interviews) is helpful so that the examiner can critically assess the interpretations from direct quotes.

Credibility can also be enhanced within the interviewing process. The reframing of questions, repetition of questions, or expansion of questions on different occasions are ways in which to increase credibility (May, 1989). Credibility is also supported when interviews or observations are internally consistent, that is, when there is a logical rationale about the same topic in the same interview or observation. In addition, indirect questions about the informants' experiences (questions such as "Do you know others who have this experience?") and structured hypothetical situations are methods that can be used to verify observations and meanings.

The credibility of any argument is enhanced by the establishment of structural coherence, that is, the assurance that there are no unexplained inconsistencies between the data and their interpretations (Guba, 1981). Although data may conflict, credibility is increased if the interpretation can explain the apparent contradictions. Accounting for rival explanations or deviant cases here is important. For example, in my disability ethnography (Krefting, 1989), informants disagreed over whether peers with brain injury were sought for friends. Rather than looking for consistency, I focused on the heterogeneity of persons with brain injury living in the community and the problems associated with grouping them together because of a shared diagnosis. As was mentioned, a range of experience or phenomena is sought in qualitative research, so that the data are not necessarily consistent but are in fact credible if described and interpreted correctly. Structural coherence is also influenced by the way that the researcher integrates in the research report the masses of loosely connected data into a logical, holistic picture.

The essence of the credibility issue is the unique authority of the researcher, the "I was there" element (Miles & Huberman, 1984). To strengthen the idea of authority, viewing the researcher as a measurement tool has been proposed. Miles and Huberman identified four characteristics that are necessary to assess the trustworthiness of the human instrument: (a) the degree of familiarity with the phenomenon and the setting under study, (b) a strong interest in conceptual or theoretical knowledge and the ability to conceptualize large amounts of qualitative data, (c) the ability to take a multidisciplinary approach, that is, to look at the subject under investigation from a number of different theoretical perspectives, and (d) good investigative skills, which are developed through literature review, course work, and experience in qualitative research methods.

One way of assessing these investigative skills or technical competence is to examine the researcher's background for any special training he or she has received that is relevant to the project, for example, experience in interviewing or observational technique. In addition, those steps that are undertaken to enhance the skills of the researcher in the specific project should be documented, for example, mock interviews, the videoing and analysis of the researcher's interviewing skills, and pilot interviews (Field & Morse, 1985). Researchers can bolster the credibility of a project by supporting their authority in these four areas.

In summary, the strategies described above are based on the concept of the researcher gathering data about, and interpreting the multiple realities of, informants. They are used to establish the truth value or credibility of the research and are critical to the accurate representation of subjective human experience. The strategies described above are not exhaustive, but represent those techniques most applicable to the types of problems studied by occupational therapists.

Transferability Strategies
As noted, there are two perspectives on applicability or transferability, and depending on one's orientation to qualitative research, transferability may or may not be an issue. If the assumption is made at the beginning of the study that the findings are descriptive in nature, representing one life perspective, as in some life histories, for example, the applicability criterion may not be relevant (Sandelowski, 1986). In such a case, data are of descriptive worth in and of themselves. If, however, the researcher means to make generalizations about the subject of the research, as is common in disability ethnography, then strategies to enhance transferability are important.

From this latter perspective, the difficulty with qualitative research is situational uniqueness; the particular group studied may not relate to others and hence conclusions may not be transferable. A key factor in the transferability of the data, then, is the representativeness of the informants for that particular group.

One strategy used to address transferability in sample selection is the use of a panel of judges to help in the selection of informants representative of the phenomenon under study. An example of this type of sample, which is also referred to as a nominated sample, is the use of one or two longtime members of a family support group to identify persons who are typical of the membership (Field & Morse, 1985). Another means of ensuring transferability is the use of a comparison of the characteristics of the informants to the demographic information available on that group being studied. As fieldwork continues, informants are selected to fill in gaps in the profile.

It is critical that researchers provide dense background information about the informants and the research context and setting to allow others to assess how transferable the findings are. As Lincoln and Guba (1985)
Dependability Strategies

Guba (1981) proposed that the dependability criterion relates to the consistency of findings. Because many qualitative methods are tailored to the research situation, there are no methodological shorthand descriptions, such as interrater reliability, commonly used in quantitative studies. The exact methods of data gathering, analysis, and interpretation in qualitative research must be described. Such dense description of methods provides information as to how repeatable the study might be or how unique the situation (Kielhofner, 1982). Guba used the term \textit{auditability} to describe the situation in which another researcher can clearly follow the decision trail used by the investigator in the study. Lincoln and Guba (1985) suggested that a single audit of the research can enhance both the dependability and confirmability of the project. This strategy, which Lincoln and Guba described in detail, will be discussed below under the confirmability criterion.

Guba (1981) also suggested that a stepwise replication technique be built into the design of a qualitative study to enhance dependability. This strategy is similar to split-half reliability in quantitative studies. Two researchers or research teams deal separately with data that have been divided, and the results are compared. An important consideration in carrying out stepwise replication is that communication between teams and team members is critical. Lincoln and Guba (1985) suggested that provision for communication on a daily basis and at preset points in the research process must be made.

Another means that the researcher can use to increase the dependability of the study is to conduct a code–recode procedure on his or her data during the analysis phase of the study. After coding a segment of data, the researcher should wait at least 2 weeks and then return and recode the same data and compare the results.

Dependability can also be enhanced through triangulation to ensure that the weaknesses of one method of data collection are compensated by the use of alternative data-gathering methods. The use of colleagues and methodological experts (peer examination) to check the research plan and implementation is another means of ensuring dependability. One can enhance stability over time by repeated observation of the same event and questioning informants about major issues; these are similar strategies to those that enhance credibility (Lincoln & Guba, 1985).

Confirmability Strategies

Guba (1981) viewed neutrality not as researcher objectivity but as data and interpretational confirmability and described the audit strategy as the major technique for establishing confirmability. This strategy involves an external auditor attempting to follow through the natural history or progression of events in a project to try to understand how and why decisions were made. In addition, auditability suggests that another researcher could arrive at comparable conclusions given the same data and research context. The auditor considers the process of research as well as the product, data, findings, interpretations, and recommendations (Lincoln & Guba, 1985).

Lincoln and Guba (1985) identified six categories of records that can be included in the audit: (a) raw data (field notes, video and audio recordings), (b) data reduction and analysis products (quantitative summaries, condensed notes, working hypotheses), (c) data reconstruction and synthesis products (thematic categories, interpretations, inferences), (d) process notes (procedures and design strategies, trustworthiness notes), (e) materials related to intentions and dispositions (study proposal, field journal), and (f) instrument development information (pilot forms, survey format, schedules). They noted that inspection and verification often are not considered until the completion of the project. In contrast, they emphasized the importance of including an auditor at the beginning of a project so that the nature of the audit trail can be determined. Ideally, the audit should be ongoing throughout the research process; the limitation in this is that the auditor could be co-opted into the project and thus lose his or her objectivity.

A number of other strategies are useful in the establishment of confirmability. Triangulation of multiple methods, data sources, and theoretical perspectives tests the strength of the researcher’s ideas. Guba (1981) noted that an investigator should provide documentation for every claim or interpretation from at least two sources to ensure that the data support the researcher’s analysis and interpretation of the findings. Another way that one can enhance neutrality is to use a team of researchers familiar with qualitative methods rather than a single researcher. Reflexive analysis is also useful to ensure that the researcher is aware of his or her influence on the data.
Summary and Conclusions

Growing interest in qualitative investigation as a legitimate approach to research questions in occupational therapy has created a need for models to assess the trustworthiness or rigor of qualitative projects. This paper has presented one such model useful both for researchers designing and conducting qualitative inquiry and for consumers of the research. Truth value, applicability, consistency, and neutrality were described as critical to the evaluation of the worth of research. These four criteria were then defined within both the quantitative and qualitative research perspectives. Several practical strategies for enhancing rigor were presented as a way for researchers to address the trustworthiness criteria.

The importance of applying such models as Guba's (1981) cannot be overstated. Occupational therapists conduct research in a climate dominated by quantitative perspectives. Grant reviewers, hospital research review committees, and journal editorial boards typically evaluate research proposals and research findings from the familiar quantitative perspective. The inclusion of a clear definition of the criteria used to assess the research and a description of how these qualitative criteria relate to quantitative criteria will help reviewers assess the value of the work. Until occupational therapists accept the principle that every qualitative research proposal and report must establish its trustworthiness, this important approach to inquiry will be considered the poor cousin of quantitative research perspectives. The knowledge base of the profession will certainly suffer without the valuable contribution of qualitative research approaches. 

Acknowledgments

I would like to acknowledge Jan Morse, PhD, for her inspirational commitment to, and scholarly mastery of, qualitative research in the health care field and the skillful editorial work of Betty Yerxa, PhD, and two anonymous reviewers.

References


