

VALIDITY AND RELIABILITY IN QUALITATIVE RESEARCH

H.I.L. Brink (Conference Paper)

Paper delivered at SA Society of Nurse Researchers' Workshop - RAU 19 March 1993

INTRODUCTION

Validity and reliability are key aspects of all research. Meticulous attention to these two aspects can make the difference between good research and poor research and can help to assure that fellow scientists accept findings as credible and trustworthy. This is particularly vital in qualitative work, where the researcher's subjectivity can so readily cloud the interpretation of the data, and where research findings are often questioned or viewed with scepticism by the scientific community.

So those of us doing qualitative studies need to be especially sensitive to the issues of validity and reliability in our projects. We need to be attuned to the multiple factors that pose risks to the validity of our findings; and plan and implement various tactics or strategies into each stage of the research project to avoid or weaken these threatening factors. We need to be aware that the tactics or strategies used to address validity and reliability in qualitative research are not the same as in quantitative research.

The very nature of qualitative research methods does not lend to statistical or empirical calculations of validity. The qualitative researcher seeks basically the same ends through different methods which are better suited to a human subject matter. A large number of authors focusing on qualitative research methods have suggested tactics or strategies the researcher can employ to enhance the truthfulness or validity of qualitative findings (Chenitz & Swanson 1986, Crabtree & Miller 1992, Field & Morse 1985, Le Compe & Goetz 1982, Morse 1991, Sandelowski 1986 and Corbin & Strauss 1990).

The purpose of this paper is to address the major risks and threats to validity and reliability in qualitative studies and in particular the tactics and strategies suggested by various qualitative researchers for avoiding or weakening the potential risks and threats. However, to refresh your memory and ensure that we all attach the same meaning to validity, reliability and qualitative research, attention will first be given to the definition and clarification of these key concepts.

THE CONCEPTS QUALITATIVE RESEARCH, VALIDITY AND RELIABILITY

The term *qualitative research* is really an umbrella term representing a variety of research approaches which share certain common elements. Qualitative researchers are not interested in causal laws but in people's belief, experience and meaning systems from the perspective of the people. Methods used are more subjective than in quantitative research and do not include statistical analysis and empirical calculation. Phenomena are viewed holistically and in their social context. Included under this umbrella term are such methods as grounded theory, phenomenology and ethnography—the three approaches earmarked for discussion today.

Validity in research is concerned with the accuracy and truthfulness of scientific findings (Le Compe and Goetz 1982: 32). A valid study should demonstrate what actually exists and a valid instrument or measure should actually measure what it is supposed to measure.

There are many types of validity and many names have been used to define the different types of validity. Campbell and Stanley (1966) have defined two major forms of validity that encompass the many types. They refer to "internal" and "external" validity, terms which are today used in most nursing research textbooks. Denzin (1970) used the distinction between internal and external validity and applied it to qualitative research. Internal validity is the term used to refer to the extent to which research findings are a true reflection or representation of reality rather than being the effects of extraneous variables. External validity addresses the degree or extent to which such representations or reflections of reality are legitimately applicable across groups.

Reliability is concerned with the consistency, *stability* and repeatability of the informant's accounts as well as the investigators' ability to collect and record information accurately (Selltiz et al 1976: 182). It refers to the ability of a research method to yield consistently the same results over repeated testing periods. In other words, it requires that a researcher using the same or comparable methods obtained the same or comparable results every time he uses the methods on the same or comparable subjects. It further requires that the researcher has developed consistent responses or habits

in using the method and scoring or rating its results and that factors related to subjects and testing procedures have been managed to reduce measurement error.

Many qualitative researchers avoid the terms validity and reliability and use terms such as credibility, trustworthiness, truth, value, applicability, consistency and confirmability, when referring to criteria for evaluating the scientific merit of qualitative research (Glaser and Strauss 1967, Leininger 1991, Lincoln & Guba 1985).

RISKS OR THREATS TO VALIDITY AND RELIABILITY

As stated previously researchers need to be attuned to the multiple factors that pose risks to the validity and reliability of their findings and plan and implement tactics or strategies to avoid or counter them. One of the key factors affecting validity and reliability is error. Error is inherent in all investigations and is inversely related to validity and reliability. The greater the degree of error the less accurate and truthful the results. Researchers thus must be especially watchful of the sources of error when planning and implementing their studies. For convenience sake the major sources of error can be categorised as follows:

- (1) the researcher
- (2) the subjects participating in the project
- (3) the situation or social context
- (4) the methods of data collection and analysis

THE RESEARCHER AS A RISK TO VALIDITY AND RELIABILITY

In a qualitative study the data-gathering instrument is frequently the researcher himself. Thus questions of researcher bias and researcher competency, if unchecked, may influence the trustworthiness of data considerably. The very presence of the researcher may affect the validity of the data provided by subjects. When a new member is introduced into an interaction reactive effects can be expected. Participants may behave abnormally (Argyris 1952). They may seek to reveal themselves in the best possible light or withhold or distort certain information; in other words the researcher has created social behaviours in others that would normally not

have occurred. Based on her extensive fieldwork, Leininger holds that researchers need to be trusted before they will be able to obtain any accurate reliable or credible data (Leininger 1991: 92).

Le Compe & Goetz refer to research findings which hold that what the researcher sees and reports is a function of the position he occupies within the participant group, the status accorded to them, and the role behaviour expected of them. The status position of the researcher can be that of an outsider or that of a participant group member. The status position can prevent the researcher from obtaining certain information and unless he is aware of this, invalid interpretation of the data may result. On the other hand, if the researcher becomes totally a part of the group there is a danger of "going native" or assuming the attitudes and behaviours of those under study. The researcher may then lose the ability to look objectively at what is happening and may develop bias towards the point of view of the group.

Researcher bias may also be introduced by the tendency of the researchers to observe subjects and interpret findings in the light of their own values, the tendency to selectively observe and record certain data at the expense of other data. The physical appearance of the researcher may influence the situation as also his dress and demeanour and personal attributes. For example, subjects may respond differently to males and females or male and female researchers may treat the subject differently.

WHAT CAN BE DONE TO DECREASE OR ELIMINATE RESEARCHER EFFECTS?

The first step in decreasing bias is to be aware of the possibility of introducing bias at various points of the research process. Field & Morse recommend that researchers undergo extensive and rigorous training as interviewers and observers before undertaking qualitative studies. Researchers need to be trained in a manner that encourages an objective view of the phenomena under study. Furthermore, every researcher should examine and declare his underlying values and assumptions in light of the research situation so that they can be considered when reading the research.

Several writers recommend that the researcher spends a period of time in the situation before data collection starts. The researcher will then become sensitised to the situation and at the same time the subjects have the opportunity to become used to the presence of the researcher (Field & Morse 1985, Le Compe & Goetz 1992, Miles & Huberman 1984).

Leininger (1991: 11) suggests that the researcher should always assess and gauge his relationships with the subjects being studied in order to enter or get close to the people or situation under study, or to move from a stranger or distrusted person to a trusted and friendly person during the research process.

She recommends the use of the stranger to field model which she has developed and validated over many years. The purpose of this model is to serve as an assessment or reflection guide for the researcher to become consciously aware of his own behaviours, feelings and responses in relation to the behaviour and experiences of subjects and as he starts to collect data, for confirmation of truths.

Once data collection starts this should be done over a long period of time. The researcher as participant observer either "lives" with the subjects as anthropologists do or spends time visiting the research site regularly over a long period of time. When subjects are interviewed over time, their responses to the same questions on the same topic should be answered with the same information. This is a type of test-retest of the same informant on the same material. The threat of "going native" or becoming so enmeshed with subjects that researchers lose their own perspective can be offset by distancing oneself from the subjects at regular intervals ie spending time away from the site, spread out site visits and discussing data with colleagues.

When field researchers are working alone, particularly when they are still unfamiliar with the setting, it is advisable that they enlist the aid of an informant who observes the occasion also. The researcher then records the activity on the spot and then reviews the written record with the informant for completeness and comprehensiveness of coverage. In some cases participant informants serve as arbiters, reviewing the days production of field notes to correct researcher misperceptions and misinterpretations. Commonly the researcher requests reactions to working analysis or processed materials from the informants. In this confirmation may be sought for various levels of the collection and analysis process (Le Compe & Goetz 1982: 42).

THE PARTICIPATING SUBJECTS AS RISKS TO VALIDITY AND RELIABILITY

The truth of responses is a key concern when data are obtained through questionnaires and interviews. Bias may be introduced because of particular responses or characteristics of the informants. Informants may want to make things seem better or worse than they are. Hospitalised patients who are questioned about the quality of their care may indicate that the care is wonderful because they fear reprisal of staff. Conversely, they may respond that their care is awful. Informants may also attempt to please the researcher by responding in the way that they believe he expects. They may also fear that by giving negative responses, they will be placed in a devalued position by the researcher. Informants may also be unwilling to share certain information with the researcher and deliberately withhold or distort it.

The researcher can attempt to increase the validity of responses in such a setting

- (1) by making sure that informants are very clear on the nature of the research eg. why the researcher is there, what he is studying, how he will collect data and what he will do with it.
- (2) by first building a trust-relationship with the subjects and staying in that setting for a long period of time
- (3) by interviewing the same informant on several occasions and making observations more than once and over time
- (4) by comparing the results obtained with other evidence
- (5) by confirming findings and analysis with informant (the danger with this technique is that subjects may become sensitised to the researcher's inferences and provide the answer that support the researcher's point)
- (6) by keeping accurate and detailed fieldnotes to note the variations in responses over the course of time
- (7) by showing fieldnotes to a second outside researcher. Another researcher is often much quicker to see where or how a fieldworker is being misled or coopted.

Informant bias may also be introduced by factors within subjects themselves such as fatigue, motivation or anxiety, duration of recall, mood, attention span, state of health and whether or not they are in pain.

To overcome this bias the researcher conducts repeated interviews at different times and in different settings and then compares results.

Another informant bias which is quoted as a particular problem in qualitative research (Miles & Huberman 1984: 230, Sandelowski 1986: 32) is the "elite bias" (overweighting data from articulate, well-informed, usually high status informants and underrepresenting data from intractable, less articulated lower-status ones).

The researcher can build in safeguards against this bias by good planning of selection of informants, by looking purposefully for contrasting cases (negative, extreme, countervailing and by carefully considering contrasting views).

THE SOCIAL CONTEXT AS A RISK TO VALIDITY AND RELIABILITY

The social context under which the data are gathered is an important consideration in establishing validity and reliability of data. Individuals may behave differently under differing social circumstances, for example, when alone with the researcher they may

provide different information than when they are in a group, or patients may provide different information within the health care context than what they reveal in their home neighbourhood. The researcher who is mindful of this will interview the same informants and make observations of behaviour in a variety of settings to make comparisons of similarities and differences before attributing meaning.

He will also specify the physical, social and interpersonal contexts within which data are gathered. Particular attention to privacy may also be of value in certain studies when subjects hesitate to answer accurately if they fear they may be overheard by others in the environment.

RISKS TO VALIDITY AND RELIABILITY PERTAINING TO DATA COLLECTION AND ANALYSIS

Because reliability and validity depend on the potential for subsequent researchers to reconstruct original strategies, the researcher who presents a vague account of his design is putting himself at risk of being accused of invalid and unreliable findings.

Researchers should therefore strive to present their methods clearly, that is precisely identify and thoroughly describe all strategies used to collect data and carefully document their field notes in the context of what was being observed to enable fellow researchers to form valid judgement. Many of the risks in data collection pertain to the researcher, respondents and social context and have already been dealt with.

Another major risk is sampling bias. The subjects under study may be overrepresenting or underrepresenting the phenomena under study. The researcher may be overreliant on accessible and elite informants. He/she may be more likely to see confirming instances of original beliefs and perceptions than to see disconfirming instances even when disconfirming instances are more frequent. From one or two concrete vivid instances he may assume that there are dozens more, but may fail to verify this (Miles & Huberman 1984: 231).

In qualitative research sample selection is based on the ability of the subject to provide data relevant to the research question. To avoid inaccurate or insufficient data, the researcher must use his/her judgement based upon the best available evidence to choose subjects who know enough, can recall enough, and are able to respond precisely to questions asked.

Secondly, the researcher should choose subjects who are able to report events not directly observable or accessible to the investigator. If the research is carried out in unfamiliar environment with an unfamiliar group of people in an unfamiliar culture, the

researcher should seek a panel of experts to assist with finding appropriate informants.

Thirdly, the researcher should do systematic "theoretical sampling" or in other words, continue to select subjects according to the findings that emerge in the course of the study. During this phase he should establish typicality or atypicality of observed events, behaviours or responses, extreme views or contrasting views. He should persist with theoretical sampling until no new information is obtained (until saturation or redundancy has been reached). There should be support for construction of a core category or several core categories which repeatedly occur while less and less new information emerges.

Another risk to representativeness of data results from the researcher's non-continuous presence. The researcher has to infer what is happening when he is not there and usually offers plausible reasons rather than evidence. To balance this risk once again multiple sources, multiple methods, multiple investigators (judge panel) varying and multiple repetitions of measurement over time are recommended.

In most qualitative approaches and particularly phenomenology, grounded theory and ethnomethods data analysis occurs simultaneously with data collection. All these methods use a series of similar steps for analysing which begin at the onset of the data collection phase.

Typical steps are coding for categories and themes and making memos about the context and variations in the phenomena under study, developing names for categories and elaborating classification systems and testing them within the data as they are collected. The findings at any point in this process will provide some direction for further data collection and the direction that the analysis may take. Judgements and inferences are made by the researcher. Major threats to the validity of data during this phase are firstly what Miles and Huberman call the "holistic fallacy" that tends to make data look more patterned or regular or congruent than they are and the tendency of the researcher to selectively observe and record certain data at the expense of other data. To avoid such selective inattention, the data analysis procedures should be exposed to a judge panel. The judge panel is selected on the basis of knowledge of content or knowledge of the research project.

Other means of providing validity and reliability are the use of the constant comparative method and the search for alternative hypothesis or negative cases (Hutchinson 1986: 116-117), checking that descriptions, explanations or theories about the data contain the typical and atypical elements of the data and obtaining validation from the subjects themselves (Sandelowski 1986: 35). Field and Morse (1989: 120) recommend that following strategies to reduce

threats to internal reliability with data analysis:-

- (1) Low inference descriptors (verbatim accounts of information provided by informants to the researcher). Use of mechanical recording enhances the accuracy of such transcripts.
- (2) Participant reviews of findings and peer examination.

CONCLUSION

In conclusion I would like to recapitulate the major critical strategies, suggested by leading qualitative researchers from various fields as essential for producing trustworthy and believable findings in qualitative research (Benner (1985), Brink & Wood (1988), Corbin & Strauss (1990), Field & Morse (1985), Glaser & Strauss (1967), Kirk & Miller (1986), Kuzel & Like (1991), Lather (1986), Le Compte & Goetz (1982), Leininger (1991), Lincoln & Guba (1985), Miles & Huberman (1984), Morse (1991), Sandelowski (1981).

1. TRIANGULATION

Triangulation refers to the use of two or more data sources, methods, investigators, theoretical perspectives and approaches to analysis in the study of a single phenomenon and then validating the congruence among them. The major goal of triangulation is to circumvent the personal biases of investigators and overcome the deficiencies intrinsic to single-investigator, single-theory, or single-method study thus increasing the validity of the study (Denzil 1989).

2. MULTIPLE repetitions of measurement over or long period of time, at different points of time, in different situations or settings and by different persons.

3. EXPERT CONSENSUAL VALIDATION FROM OTHERS familiar with the topic under study at various stages of the research process. This involves independent analysis of the data by others (eg research colleagues, a judge panel or participant informants. (For example, following each initial category generation during analysis - independent fellow researchers or other experts in the field are invited to generate their own category system. This is then compared with the researchers).

4. MEMBER CHECKS

This refers to recycling of analysis back to informants. (Conclusions are given to the informants and feedback is requested about the accuracy of the content. This ensures that the researcher and the informant are viewing the data consistently.

5. SEARCHING FOR DISCONFIRMING EVIDENCE

In this step there is an active search for disconfirmation of what is believed to be true. A proposition deserves some degree of trust only when it has survived serious attempts to

falsify it. (Cronbach in Lather 1986: 67). Searching for disconfirming evidence involves both purposive sampling and prolonged engagement with informants in the field. Purposeful sampling allows the researcher to include informants who may differ from key informants in critical ways. This purposeful sampling of individuals and the inclusion of conflicting as well as complementary accounts strengthens the description. One varies or contrasts the conditions as methodically as possible in order to determine what has an impact on the phenomenon in question (Corbin & Strauss 1990).

6. CHECKING FOR REPRESENTATIVENESS

This step includes checking for the representativeness of the data as a whole, of the coding categories and of the examples used to analyze and present the data.

7. THICK DESCRIPTION

Qualitative research, like its quantitative cousins, can be systematically evaluated only if their criteria and procedures are made explicit. Hence validity and reliability can only be judged if a very detailed account of the context or setting within which the study took place and a thorough description of the procedures from the beginning to the end is given. Most qualitative researchers refer to this as thick description. However Lincoln and Guba (1985) use the term auditability. Auditability means that any reader or another researcher can follow the progression of events in the study and understand their logic. Such an account starts with the researcher recognising and disclosing initial assumptions, suppositions and values that may have influenced data gathering and processing. The researcher furthermore describes, explains and justifies.

- (1) how he/she becomes interested in the subject matter of the study
- (2) the specific purpose(s) of the study
- (3) how subjects or pieces of evidence came to be included in the study and how they were approached, how theoretical sampling was done
- (4) the impact, the subjects and the researcher/s had on each other
- (5) how the data were collected ie
 - the number and types of data collection methods and procedures used in the study
 - the number of times data are collected on the same subject using the same method
 - how often and over what period were data collected on the same and different informants
 - the number of investigators collecting the same and different information on the same and different subjects; and

- (6) when and how data were analysed
- (7) the nature of the setting(s) in which data were collected
- (8) how various elements of the data were weighted
- (9) the inclusiveness and exclusiveness of the categories developed to con the data and
- (10) how reliability and validity were answered

The account also needs to include the relationship between field notes and the conclusions based upon them. Memos, data display charts, indicating coding instructions and the actual placement of data into categories and the way different elements of the data were linked to each other, should be kept available.

Corbin and Strauss (1990:20) strongly urge investigators to adhere to the major criteria for answering credibility unless there are exceptional reasons for not doing so. In such unusual cases, researchers should know precisely how and why they depart from the criteria, say so in their writing, and submit the credibility of their findings to the reader. If we want to develop a valid, truthful and believable account of our qualitative studies it will be wise to take note of and implement the critical strategies discussed unless otherwise indicated.

LIST OF SOURCES

- Argyris, C. 1952. Diagnosing differences against the outsider. *Journal of Social Issues* 8 (no 3): 24-34.
- Benner, P. 1985. Quality of life: A phenomenological perspective. *Advances in Nursing Science* 8 (no 1): 1-14.
- Campbell, D.T. & Stanley, J.C. 1963. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Chenitz, W.C. & Swanson, J.M. 1986. *From practice to grounded theory*. Menlo Park CA: Addison Wesley.
- Corbin, J. & Strauss A. 1990. Grounded theory research: Procedure, canons and evaluative criteria. *Qualitative Sociology* 13 (no 1): 3-21.
- Crabtree, B.F. & Miller, W. 1992. *Doing qualitative research*. Newbury Park, CA: Sage.
- Denzin, N.K. 1970. *The research act: A theoretical introduction to sociological methods*. Chicago: Aldine Publishing Co.
- Field, P.A. & Morse, J.M. 1985. *Nursing Research. The application of qualitative*

approaches. Beckenham Kent: Croom Helm.

Glaser, B. * Strauss A. 1967. *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Publishing Co.

Kirk, J. & Miller M. 1986. *Reliability, validity and qualitative research*. Beverly Hills CA: Sage.

Kuzel, A.J. & Like, R.C. 1991. Standards of trustworthiness for qualitative studies in primary care. In P. Norton, M. Stewart, F. Dudiver, M. Bass & E Dunn (eds). *Primary Care Research: Traditional and innovative approaches*. (p.138-158). Newbury park CA: Sage.

Lather, P. 1986. Issues of validity in openly ideological research: Between a rock and a soft place. *Interchange* 17: 63-84.

Le Compe, M.D. & Goetz, J.P. 1982. Problems of reliability and validity in ethnographic research. *Review of Educational Research* 52 (no 1): 31-60.

Leininger, M. 1991. *Culture care, diversity and universality: A theory of nursing*. New York: N.L.N.

Lincoln, Y. & Guba, G. 1985. *Naturalistic inquiry*. Beverly Hills: Sage.

Miles, M.B. & Huberman, A.M. 1984. *Qualitative data analysis: A sourcebook of new methods*. Beverly Hills CA: Sage.

Morse, j.M. 1991. *Qualitative nursing research: A contemporary dialogue*. Rockville, Maryland: Aspen.

Oiler, C. 1982. The phenomenological approach in nursing research. *Nursing Research* 31 (no 30): 178-181.

Reason, P. & Rowan, J. 1981. *human inquiry: A sourcebook of new paradigm research*. New York: John Wiley & Sons.

Sandelowski, M. 1986. The problem of rigor in qualitative research. *Advances in Nursing Science* 8 (no 3): 27-37.

Seltiz, C. & Wrightsman, L.C. & Cook, W.S. 1976. *Research methods in social relations*. 3rd edition. New York: Holt Rinehart & Winston.

Strauss, A. & Corbin, J. 1990. *Basics of qualitative research*. Newbury Park CA: Sage.

H. Brink
Professor, Department of Nursing Science
UNISA