VALIDITY IN QUALITATIVE RESEARCH

Janet N. Rosenbaum

The complexity of human experience requires a variety of research methods for understanding nursing phenomena. In addition to quantitative methods, qualitative research, which brings a different research tradition to nursing, is gaining interest amongst researchers. However, it is time to move beyond debating the value of qualitative or quantitative research in nursing (Goodwin & Goodwin, 1984; Munhall, 1982.) Assuming that different methods help answer different research questions (Polit & Hungler, 1987; Leininger, 1985), it is timely to demonstrate the rigour of qualitative research.

The purpose of this paper is to discuss the issue of validity in qualitative research. Relatively little space is devoted explicitly to the topic of validity in many qualitative research textbooks (Munhall & Oiler, 1986, Parse, Coyne & Smith 1985) or in research reports giving readers the false impression that validity concepts receive only minor consideration. This author believes that qualitative researchers must increase documentation of the rigour of their research in order to gain credibility within the nursing science community.

The following questions will be answered in this paper.
1. Is instrument validity relevant in qualitative research?
2. Are concerns about internal and external validity relevant in qualitative research?
3. What are the threats to internal and external validity in qualitative research?
4. What attention is given to validity in qualitative nursing research textbooks and research reports?

Definitions of Validity

The literature identifies validity in a number of ways. One is in relation to instruments and the purpose for which they are used (Carmines & Zeller, 1979; Kerlinger, 1986; Nunnally, 1978). Cook and Campbell (1979) view validity differently, from the perspective of truth value of propositions rather than measurement issues in instrumentation. This paper considers both views of validity.

Janet N. Rosenbaum, RN, M.Sc.N., is Associate Professor in the School of Nursing at the University of Windsor, Windsor, Ontario and a Ph.D. student in Nursing at Wayne State University, Detroit, Mich.

Nursing Papers/Perspectives en nursing 1988, 19(3), 55-66

Validity (instrument)

"Validity refers to the degree to which an instrument measures what it is supposed to be measuring" (Polit & Hungler, 1987, p. 323). It is guided by the question, "Are we measuring what we think we are measuring?" (Kerlinger, 1986, p.417). Further, there is the presupposition that reliability is necessary for validity (Polit & Hungler, 1987).

Internal validity

"Internal validity refers to the approximate validity with which we infer that a relationship between two variables is causal or that the absence of a relationship implies the absence of cause" (Cook & Campbell, 1979, p. 37).

External validity

"External validity refers to the approximate validity with which we can infer that the presumed causal relationship can be generalized to and across alternate measures of the cause and effect and across different types of persons, settings, and times" (Cook & Campbell, 1979, p. 37).

Validity (Instrument)

Validity, as it pertains to instrument measures, is sometimes viewed as having limited appropriateness in the discussion of qualitative research. According to Nunnally's (1978, p. 3) definition of measurement as "rules for assigning numbers to objects in such a way as to represent quantities of attributes," qualitative research does not use measurement. However, Kerlinger (1986, p. 391) says that "in its broadest sense, measurement is the assignment of numerals to objects or events according to rules", allowing for nominal level of measurement. In qualitative research, recurrent features such as patterns and themes are identified using classification systems (Leininger, 1985). Atwood and Hinds (1986) make a case that measurement is used in the grounded theory method and note that little attention has been given to validity and reliability.

Further, a number of qualitative research sources refer to the researcher as instrument (Dobbert, 1982; Lincoln & Guba, 1985). Dobbert says that the researcher, as instrument, may bring biases to data gathering. Hence a general discussion of validity (instrument), with the focus on content validity as being most relevant, proceeds before internal and external validity are explored. Researchers must be aware, through every phase of research, of their own likes, dislikes and prejudices as they relate to the participants or group being studied. Dobbert (1982) stresses that it is the unrecognized biases that cause validity problems; therefore, the researcher, as instrument must indulge in careful self-examination.

Lincoln and Guba (1985, p. 193) list the following benefits of the human as instrument:

1. Responsiveness
2. Adaptability
3. Holistic emphasis
4. Knowledge base expansion (domains of propositional and tacit knowledge)
5. Processual immediacy (process data as soon as available)
6. Opportunities for clarification and summarization
7. Opportunities to explore atypical or idiosyncratic responses

However, these authors say that human instruments, as well as paper and pencil instruments, are fallible and value laden. They suggest multiple benefits of using teams of "human instrumentation" (p. 236) to improve validity. For example, a team is more likely to have a member who shares the values of the informants than is one researcher alone. Teams may represent multiple disciplines, strategies and methods. They may provide checks of validity through mutual consultations, and sharing insights. With careful team training and implementation of strategies, the researchers as instruments are able to provide credibility.

In summary, instrument validity is a relevant factor in qualitative research when viewing the researcher as instrument, and when searching for themes and patterns.

Internal and External Validity

Cook and Campbell (1979, p. 37) consider validity in experimental research as "the best available approximation to the truth" propositions. They use the word "approximation" because they say that truth is never certain.

Are concerns about internal and external validity relevant in qualitative research? Le Compte and Goetz (1982) discuss the issue of validity from an ethnographic research perspective. They acknowledge that ethnographers historically have been less concerned about validity than quantitative researchers and hence have been subject to criticism. However, these authors identify several major differences between quantitative and ethnographic research that result in differences in formation of research problems, nature of goals and application of results of research.

Formation of the research problem in quantitative research involves eliminating extraneous and contextual factors in order to examine effects caused by a treatment. However, in ethnographic research the context is focal to inquiry

that rarely involves a treatment. "The naturalistic setting in which ethnography normally is conducted both facilitates on-the-spot analysis of causes and processes and precludes precise control of so-called extraneous factors" (Le Compte & Goetz, 1982, p. 33).

The nature of goals in qualitative and quantitative research differs especially regarding the stage of the research when theory is prominent (Le Compte & Goetz, 1982). Ethnographers often avoid a priori hypotheses or propositions using an inductive theory generating approach. In contrast, quantitative research is concerned with theory testing.

Generalization is often intended in quantitative research since the ideal is to have subjects who are randomly sampled (Le Compte & Goetz, 1982). Random sampling is often unrealized; therefore such strategies as replication and encouraging co-operation in recruitment of subjects to have representative subjects in a variety of settings may increase generalizability (Cook & Campbell, 1979). However, Le Compte and Goetz (1982) say that ethnographers usually do not have the opportunity to use these same strategies. Therefore they say that terms that are more appropriate for use in ethnography instead of generalization are comparability and translatability. Describing the characteristics of the group being studied permits decision about comparison with other groups (p. 34). By clearly identifying research methods and categories for analysis, translatability combines with comparability to provide a qualitative analog to generalizability across groups.

Le Compte and Goetz (1982) claim the high internal validity of ethnographic research is attributable to the data collection and analysis techniques. Pelto and Pelto (1978) show that living amongst the participants, collecting data over long periods of time using ethnographic interviewing techniques, doing on going data analysis, using natural settings that are reality for the participants and researcher self-monitoring are among the strategies that these authors say contribute to internal validity.

Le Compte and Goetz (1982) acknowledge that Cook and Campbell's (1979) threats to validity are relevant to ethnographic research and should be addressed even though the problems present themselves differently. Specifically, they identify the following threats to internal validity in ethnographic research: history and maturation, observer effects, selection and regression, mortality and spurious conclusions.

History is acknowledged to effect data collection (Le Compte & Goetz, 1982, p. 45) and "the ethnographic task is to establish which baseline data remain stable over time and which data change." This is done by establishing long-term residence in the field, using time-sampling procedures, comparing baseline data, revisiting sites and replicating. Many of these same strategies

also control for maturation. However, "ethnographers view maturational stages as varying according to cultural norms" (Le Compte & Goetz, 1982, p. 45). Therefore to control for maturation, ethnographers take care to differentiate behaviors which are expected in different socio-cultural groups.

What Le Compte and Goetz (1982) identify as observer effects are analogous to Cook and Campbell's (1979) testing and instrumentation threats in quantitative research. They acknowledge that participant observation, informant interviewing and auxiliary use of instruments all pose problems related to internal validity. Effects of observers' presence as a threat is compensated by establishing a number of field relationships, maintaining an appropriate balance of engagement and neutrality in informant relationships, analyzing field relationships and remaining in the field long enough to reduce factitious informant responses. "Going native", a problem that occurs when the researcher is over-involved to the point that objectivity can no longer be maintained, requires temporary disengagement (Le Compte & Goetz, 1982, p. 7).

Researchers should be aware of their own ethnocentric bias in guarding against observer effects. Le Compte and Goetz (1982) say that researchers must ensure that the perspective of the informants is maintained by getting feedback from the informants to affirm that the researchers have made correct interpretations.

Selection and regression effects which are of concern to ethnographic researchers, are tempered in a number of ways. By compiling an inventory of sub-groups in the field, interacting with a diversity of informants, and questioning assumed meanings, representative findings are more likely to result (Le Compte & Goetz, 1982). The issue of mortality is also addressed by these authors because groups change over time. By collecting baseline data, the losses and gains in group memberships will be documented, and their effects will be identified. However, "the naturalistic approach prevents the interchangeability of human informants and participants" (p.49).

As for external validity, Le Compte and Goetz (1982) say that the nature of ethnographic research often prevents generalizing across groups. Descriptive data are usually sought regarding relatively unknown phenomena for the purpose of explication, and informants are not randomly selected. These authors include the following threats to external validity that reduce cross-group comparisons in ethnographic research: selection, setting, and history effects.

Selection effects are a problem when the researcher designates categories which do not match the reality of the group studied. "Some constructs cannot be compared across groups because they are specific to a single group or

because the researcher mistakenly has chosen groups for which the construct does not obtain" (Le Compte & Goetz, 1982, p. 51). Denzin (1978) says similarly that participant observers must study informants who are representative of their group, if generalizations are made. Setting effects may occur by the very act of studying a group, hence constructs may not be comparable across settings because they may be context specific. "When the construct is a function of observer-setting interaction, it may be treated as equivalent only for groups being observed in a comparable manner, and the interactive dynamics should be identified clearly" (Le Compte & Goetz, p. 52). As well, some settings become "oversaturated" because they are frequently the source of research studies and are likely different from settings that have not received research attention. History is recognized as an external validity threat because cross-site comparisons are ineffective when unique experiences make the groups quite different.

Dobbert (1982) reminds us that validity is relative to the intended goals of a particular research project, the method and instrumentation. In ethnographic research, the primary goal is to discover patterns while understanding the situation from the participants' perspective. "There is always more than one view of any social situation" (p. 261), and human situations are ever changing. He states that three factors influence the perspective of informants: social position in the group, personality, and status and role. Dobbert emphasizes the need for multiple methods to increase validity because more perspectives permit cross-checking of data. For instance, using methods providing permanent data (such as tapes), in addition to interviews, attending cultural ceremonies such as wedding and funerals, and matching data against data collected in similar contexts are some ways to improve validity. Denzin (1978) concurs with the notion of multiple data sources by saying that the "greatest weight is placed on those hypotheses which have withstood the impact of triangulated observations" (p. 107). Fielding and Fielding (1986) support triangulation by the linkage of diverse data, researchers and techniques in order to improve validity.

Lincoln and Guba (1985) bring an interesting perspective to the discussion of validity in naturalistic inquiry. They say that research within the naturalistic paradigm has 14 components that build upon the naturalist axioms in Table I (pp. 39-43).

1. Natural setting.
2. Human instrument (the researcher and respondent).
3. Utilization of tacit knowledge (to learn nuances of multiple realities).
4. Qualitative methods (although not exclusively).
5. Purposive sampling (to uncover the full array of multiple realities).
6. Inductive data analysis.
7. Grounded theory (rather than a priori theory).

8. Emergent design (to allow research design to emerge.
9. Negotiated outcomes (respondents can best understand and interpret the influence of local value patterns).
10. Case study reporting mode.
11. Idiographic interpretation (in terms of particulars of the case rather than nomothetically in terms of lawlike generalizations).
12. Tentative application (of findings due to multiple realities).
13. Focus-determined boundaries (multiple realities define focus of research rather than inquirer preconceptions).
14. Special criteria for trustworthiness (internal and external validity, reliability, and objectivity).

Lincoln and Guba (1985) say that the usual criteria for internal validity fail because of the multiple realities of naturalistic inquiry, necessitating alternative criteria. Instead, in order to demonstrate "truth value", "credibility" is demonstrated by approval of the constructors of the multiple realities being studied - the informants (p. 296). These authors contrasts threats to internal validity between naturalistic inquiry and quasi-experimental research. They say, "since changes do occur in human instruments" (p. 296), instrumentation is a greater threat to credibility in naturalistic settings. They also say that another threat, statistical regression, does not apply to naturalistic settings unless quantitative tools are used. Finally, maturation and maturation/selection interaction are less threatening in naturalistic inquiry than quasi-experimental research because of long-term interactions between researcher and informants.

Table 1

Axioms of the Naturalist Paradigm

Axioms About:	Naturalist Paradigm
Nature of reality	Realities are multiple
Relationship of knower known	Knower and known are interactive and to inseperable
Generalization	Only time-and context-bound working hypothesis (idiographic statements) are posible
Causal linkages	Entities are in a state of mutual shaping – imposible to distiguish causes from effects
Role of values	Inquiry is value-bound

Adapted from: Lincon, Y. & Guba, E. (1985). *Naturalistic Inquiry*. Beverly Hills: Sage Publications, p. 37.

Lincoln and Guba (1985) write about transferability as opposed to external validity. At the heart of transferability is contextual similarity - congruence between contexts. When the original investigator provides "thick description" (p. 359), that is, a comprehensive base of information, judgement about transferability can be made. However, in many cases, data collected in a particular context has meaning only for that context at that particular time. By using purposeful sampling, as much information as possible will be included in the "thick description" by the naturalistic researcher to provide data for transferability determination.

Qualitative Methodology

The following section is divided to include perusal of qualitative nursing research textbooks and research reports to view nursing authors' attention to validity.

Qualitative nursing research textbooks

There has been a rising interest in qualitative research in nursing that is reflected in the publication of five textbooks since 1985 (Chenitz & Swanson, 1986; Field & Morse, 1985; Leininger, 1985; Munhall & Oiler, 1986; Parse, Coyne & Smith, 1985) with variable attention to validity.

The focus of Chenitz and Swanson (1986) is grounded theory. These authors give extensive coverage to Cook and Campbell's (1979) concerns about internal and external validity. Even though they state that validity is critical for evaluating research, they suggest that credibility is a better word to use than validity.

Parse et at. (1985) discuss many of the concepts which other authors say lead to validity without naming validity directly. Instead, when they write about evaluation of qualitative research, they name several standards which may be related to validity: substance, clarity and integration. They speak of "validating" (p. 69) as contributing to knowledge of a particular culture using the ethnographic method.

In the textbook by Munhall and Oiler (1986), principles and techniques of phenomenology are discussed. While validity is not named, as in Parse et al. (1985), concepts related to validity are explored. "Attention to subjects' realities", "data gathering procedures...preserve the natural spontaneity of subjects' lived experiences" and "the researcher must recognize that he or she is immersed in the phenomenon of study by virtue of studying it" are examples of attention to truth (p. 80). Rieman (1986), in a chapter of this book speaks of validating themes by referring to original descriptions and by asking the respondents if they agree with the descriptions.

Leininger (1985) addresses validity directly, even though she emphasizes that the same criteria should not be used in qualitative research as in quantitative research. She says that criterion-related validity can be obtained in qualitative research with "congruency, meaning, and syntactical relationships of findings" based upon "knowing and understanding the phenomena" (p. 69). She also states that predictive validity allows prediction of "human lifeways or behaviour" (p. 69). She clearly states the goal of qualitative validity.

> Measurement is not the goal; rather, knowing and understanding the phenomena is the goal. Qualitative validity is concerned with confirming the truth or understandings associated with phenomena (p. 68).

Finally, Field and Morse (1985) devote considerable space to their discussion of validity which they relate to Le Compte and Goetz's (1982) article. They define validity in qualitative research as "the extent to which the research findings represent reality" (p. 139). These authors acknowledge the need to attend to validity and review many threats to internal and external validity.

Qualitative nursing research reports

A convenience sample of nine qualitative nursing research reports was chosen, representing ethnography/ethnonursing, grounded theory, and phenomenological studies (Aamodt, 1986; Benoliel, 1983; Boyle, 1984; Drew, 1986; Kus, 1986; Lynch-Sauer, 1985; Moore, 1984; Rieman, 1986; and Stern, 1980). These studies were examined to see if concepts related to validity to the discussion of validity (or its equivalent) as in studies on quantitative research.

All nine studies addressed concepts related to validity. For instance, Stern (1980) discussed selective sampling in her grounded theory study to expand categories until saturation occurs. Drew (1986) referred to the use of the constant comparative method of analysis in phenomenology which compares the content of interviews to previous interviews. Aamodt (1986) talked about repeated interviews of children in her ethnographic/ethnonursing study. For the most part, validity concepts were integrated throughout the research reports rather than discussing them separately in one section. Only Drew's study (1986) devoted a separate section to validity, which also included reliability.

Summary and Conclusion

This paper discusses various types of validity and threats to validity in qualitative research. Additionally, qualitative nursing textbooks and research reports were examined for attention to concepts related to validity. This writer concludes that issues relative to validity are important in qualitative nursing research even though the research problems, goals and application of results are

different than quantitative research. The review of qualitative nursing textbooks and research studies leads this writer to conclude that widespread attention is given to validity concepts. However, the research consumer needs more explicit documentation of this validity effort. Separate sections in research reports, and entire chapters in qualitative nursing research textbooks will document attention to validity and the rigour of the research.

REFERENCES

Aamodt, A. (1986). Discovering the child's view of alopecia: Doing ethnography. In P. Munhall, & C. Oiler, *Nursing Research: A qualitative perspective* (pp. 163-172). Norwalk: Appleton-Century-Crofts.

Atwood, J., & Hinds, P. (1986). Heuristic heresy: Application of reliability and validity criteria to products of grounded theory. *Western Journal of Nursing Research*, 8, 135-154.

Benoliel, J. (1983). Grounded theory and qualitative data: the socializing influences in life-threatening disease on identity development. In P. Wooldridge, M. Schmitt, J. Skipper, & R. Leonard (Ed.), *Behavioral science & nursing theory* (pp. 141-187) St. Louis: C.V. Mosby.

Boyle, J. (1984). Indigenous caring practices in Guatemalan Colonia. In N. Leininger (Ed.), *Care: The essence of nursing and health* (pp. 123-132 Thorofare: Slack Inc.

Carmines, E., & Zeller, R. (1979). *Reliability and validity assessment.* Beverly Hills: Sage Publications.

Chenitz, W., & Swanson, J. (1986). *From practice to grounded theory: Qualitative research in nursing.* Menlo Park: Addison-Wesley Publication.

Cook, T., & Campbell, D. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin.

Denzin, N. (1978). *The research act: A theoretical introduction to sociological methods.* New York: McGraw-Hill Book.

Dobbert, M.L. (1982). *Ethnographic research.* New York: Praeger Publishers.

Drew, N. (1986). Exclusion and confirmation: A phenomenology of patient's experience with caregivers. *Image*, 18(2), 39-43.

Field, P.A., & Morse, J. (1985). *Nursing research: The application of qualitative approaches.* Rockville: Aspen Systems Corporation.

Fielding, N., & Fielding, J. (1986). *Linking data.* Sage University Paper series on Qualitative Research Methods (4). Beverly Hills: Sage.

Goodwin, L., & Goodwin, W. (1984). Qualitative vs. quantitative research or qualitative and quantitative research? *Nursing Research*, 33, 378-380.

Kerlinger, F. (1986). *Foundations of behavioral research.* New York: Holt, Rinehart and Winston.

Kus, R. (1986). From grounded theory to clinical practice: Cases from gay studies research. In W. Chenitz, & J. Swanson, *From practice to grounded theory* (pp. 227-240). Menlo Park: Addison-Wesley Publishing.

Le Compte, M., & Goetz, J. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52, 31-60.

Leininger, M. (1985). *Qualitative research methods in nursing.* Orlando: Grune & Stratton.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry.* Beverly Hills: Sage Publications.

Lynch-Sauer, J. (1985). Using phenomenological research method to study nursing phenomena. In M. Leininger (Ed.), *Qualitative research methods in nursing* (pp. 93-107). Orlando: Grune & Stratton, Inc.

Moore, M. (1984). Self-care and caretaking of the adolescent asthmatic girl. In M. Leininger (Ed.), *Care: The essence of nursing and health* (pp. 123-132). Thorofare: Slack Inc.

Munhall, P. (1982). Nursing philosophy and nursing research: In apposition or opposition? *Nursing Research*, 31, 176-177, 181.

Munhall, P., & Oiler, C. (1986). *Nursing research: A qualitative perspective.* Norwalk: Appleton-Century-Crofts.

Nunnally, J. (1978). *Psychometric theory.* New York: McGraw-Hill Co.

Parse, R., Coyne, A., & Smith M. (1985). *Nursing research: Qualitative methods.* Bowie: Brady Communications Company.

Pelto, P., & Pelto, G. (1978). *Anthropological research: The structure of inquiry.* Cambridge: Cambridge University Press.

Polit, D., & Hungler, B. (1987). *Nursing research: Principles and methods.* Philadelphia: J.B. Lippincott.

Riemen, D. (1986). The essential structure of a caring interaction: Doing phenomenology. In P. Munhall, & C. Oiler, *Nursing Research: A qualitative perspective* (pp. 85-108). Norwalk: Appleton-Century-Crofts.

Stern, P. (1980). Grounded theory methodology: Its uses and processes. *Image*, 12(1), 20-23.

The author wishes to acknowledge Dr. Judith Floyd, Wayne State University College of Nursing, for her encouragement and for reading an earlier version of this paper.

RÉSUMÉ

Recherche qualitative et validité

Afin d'acquérir une plus grande crédibilité dans les milieux infirmiers, des chercheurs spécialisés en recherche qualitative doivent justifier la rigueur de leur recherche. A cette fin, ils peuvent faire état des mesures prises pour en garantir la validité. Bien que les problèmes, les objectifs et l'application des résultats de la recherche ne sont pas les mêmes en recherche qualitative et en recherche quantitative, la validité demeure importante dans les deux cas.