Theory-testing research: Building the science

Lack of emphasis on the empirical validation of nursing models has hindered the development of nursing science. This article explores issues concerning theory development in nursing and criteria for evaluating theory-testing research. From a comprehensive review of the literature on theory testing, 15 criteria are identified as essential for evaluating theory-testing studies. The criteria are applied to selected research reports, and recommendations for future use of criteria are suggested.

Gayle J. Acton, RN, MSN
Doctoral Student

Barbara L. Irvin, RN, MS, MN
Doctoral Student
School of Nursing
University of Texas at Austin
Austin, Texas

Barbara A. Hopkins, RN, MSN, CNAA
Doctoral Student
School of Nursing
University of Texas at Austin
Austin, Texas
Nurse Educator
Nursing Education and Research Department
Scott & White Memorial Hospital
Temple, Texas

THE PAUCITY OF REPORTED research that specifically tests nursing models is a deterrent to the development of the scientific base for nursing. Further, the lack of attention to criteria for evaluating theory-testing research has led to the inadequate use of nursing models for guiding the design and analysis of many nursing studies. This article will address problems related to theory-testing research and offer criteria that the authors hope will prove useful for advancing the science of nursing.

BACKGROUND OF THE PROBLEM

To establish nursing as distinct and separate from other health care professions, we nurse scientists must identify and validate that body of knowledge that is uniquely ours. Essential to this task is the establishment of a solid theoretic base upon which to develop knowledge, conduct research, and guide practice. The lack of such a base has been

Adv Nurs Sci 1991;14(1):52–61

identified as an important problem facing nursing today.[1] To overcome this problem, support for theory in the form of testing must be demonstrated.[2]

Nursing is in the early stages of developing its science. Typical of a young science are theoretic gaps and inconsistencies. These inconsistencies are usually the result of newly discovered concepts and relationships that do not fit into or are not addressed by existing theories, thereby creating theoretic gaps.[3] Such gaps exist in nursing today, although theory development has been a goal since the 1950s. It has only been in the last decade that nursing has been able to agree on a universal definition of nursing. With the publication of the *ANA Social Policy Statement*,[4] the focus of nursing changed from the disease process to the identification of human responses to health and illness. This focus created new concepts and relationships to be discovered and explored. Testing of both new and established concepts and relationships, however, has not earned priority status.

Nurse leaders have called for increased theory validation for the past 30 years. Evidence that unique theories are being developed is provided by Moody,[5] who compiled a list of 11 established and 30 emerging nurse theorists. However, few authors responded in 1984 when Chinn,[6] as editor of *ANS*, called for manuscripts reporting theory-testing research. Further, Moody's analysis of nursing research from 1977 to 1986[7] showed that, of 720 studies described in six research journals, only 3% tested theory, concepts, or hypotheses from an explicit conceptual or theoretical framework. Silva[8] also reported that only 9 of 62 research reports identifying a specific nursing theory fell into the category of theory testing. One can conclude from these analyses that the nursing profession has been more concerned with the development of theory than with its testing. Yet theories for nursing are only as strong as their tests of accuracy.[9,10]

Theory-testing research seeks to develop evidence about hypotheses derived from a theory. Thus, deductive reasoning forms the basis for theory testing. Deductive reasoning begins with broad theoretic concepts and narrows to specific observations. That is, it explains and predicts what will be observed, thus converting theory into relational statements, propositions, and hypotheses that can be tested by appropriate research methods.[11]

The question "Can nursing theory be tested?" was both raised and answered by Jacobs, who responded, "Yes and no, perhaps."[12(p39)] This apparent ambivalence results from the problem of directly testing nursing theory. Tests of the grand theories are problematic, because the abstract theoretical constructs do not lend themselves to empirical measurement. Regardless of this dilemma, however, Fawcett[13] stressed that nursing must continue to work on theory development. She emphasizes the need to create conceptual–theoretical–empirical structures that will allow for theory testing. Such structures identify and describe nursing conceptual model constructs, including propositions, that result in theory level concepts and propositions that can be empirically measured and tested.

Despite the emphasis on testing, theories cannot be proven or disproven. Theories are dynamic and, at best, tentative. They attempt to explain what has been observed in reality and to predict what will be observed in research.[11] Theory-testing research evaluates how well the explanations and predictions that evolve from the theory hold up and thus

refines theory into a clearer and more useful depiction of reality.[9,14]

Through scientific research, a theory may be validated, refuted, or modified. Typically, several studies are necessary to test a theory adequately, and the more rigorous the test, the more valuable the theory becomes. Therefore, it is only when theories are tested and refined repeatedly that their worth can be evaluated.[10,15]

REVIEW OF LITERATURE

To identify criteria for evaluating theory-testing research, a comprehensive search of the nursing literature, as well as a review of theory-testing literature in the social sciences, was conducted. The articles reviewed were primarily those published between 1986 and 1990. Potential sources were obtained by searching titles, abstracts, and keyword indices using the terms theory, testing, and research. Studies for this article were not restricted to any particular nursing model, unlike the search by Silva[8] that was limited to five nursing models. In fact, an effort was made by the authors to include new nursing models that are examples of developing mid-range theories.

In her now classic article, Silva[8] addressed the pressing need for the development and testing of nursing theory to establish the scientific base for nursing. She pointed out the lack of clarity about what constitutes theory testing, particularly with regard to nursing theory. Silva identified the degree to which empirical research explicitly tested nursing theory by using specified evaluation criteria. She evaluated and reported three exemplary cases of empirical research that tested theoretic constructs or propositions derived from a nursing grand theory. In addition, Silva discussed various implications for nursing theory, practice, and research, including factors that impede and enhance the testing of nursing theory.

One of the major flaws identified by Silva[8] is that many researchers have mistakenly labeled their work as theory-testing research. In fact, most authors merely use an identified model or theoretic framework for organization of the study or structuring of instruments. Because of this Silva classified the 62 studies she analyzed into three categories: minimal, insufficient, and adequate use of models to constitute theory testing. Of these studies, 24 fell into the minimal category, 29 into the insufficient category, and only 9 into the adequate category.

As a result of these findings, Silva[8] determined that an expressed purpose of a theory-testing study is to examine the underlying validity of the assumptions or propositions of a selected nursing model. She also stated that a nursing model must be explicitly identified as a theoretic framework for the research. The model must be discussed comprehensively enough for the connection of the hypotheses or purposes and the model to be clear.

Other criteria identified by Silva[8] are that the hypotheses or purposes must be empirically tested in an appropriate manner, and that, as a result of this test, indirect evidence exists of the validity (or lack thereof) of the designated assumptions or propositions. Silva also stated that authors should present a discussion of the extent to which the evidence explains, supports, or refutes the relevant aspects of the model or theory.

Walker and Avant[16] adapted Silva's criteria to make them more specific and concise. They also added the following criterion: "The hypotheses used to test a specific theory are designed to put the theory at risk for falsification by virtue of their specificity

and compatibility with only a limited set of events."[16(p199)] These authors argued that "the more specific the predictions that can be made from a theory, the more readily it can be falsified and the narrower the range of data that will support the theory."[16(p199)] They further stated that a theory can be judged as valid and falsifiable when the results of testing highly specific hypotheses yield data that closely fit predictions.

In a comparative discussion of theory-testing versus theory-generating research, Chinn and Jacobs[14] articulated theory-testing criteria. These authors asserted not only that the portion of the theoretic framework under study must be described explicitly and summarized, but also that the literature review must describe previous studies based on the theory and must trace how the study has been conceived. The research problem may be stated in question form, but it must clarify how the research is to be accomplished, clarify the variables or events to be studied, and reflect possibilities for operationalizing abstract concepts.[14]

Chinn and Jacobs[14] also discussed the importance of the theoretic adequacy of instruments and approaches to assessing changes that can be expected on the basis of the theory. "The problems of reliability and validity of both direct and indirect observations . . . are also important in theory-testing research."[14(p161)] Further, these authors state that sample selection should be such that subject responses will support or refute the theory. Obviously, this statement assumes that the requirements for statistical analysis will be met.

Chinn and Jacobs[14] asserted that testing relationships between variables requires deliberate manipulation of circumstances so that what was hypothesized can be assessed. Accordingly, the design must be consistent with the theoretic basis for the study. In addition, data analysis "must be consistent with the purposes of the research as well as the research design."[14(p163)] The analysis should yield enough quantitative and qualitative evidence to answer the research question or to support or reject the hypotheses. An interpretive analysis of the results as they pertain to the theory should be included in the study conclusions. In addition, the conclusions should address the theoretic significance of the research.

To supplement the above discussion, it is worth noting a few observations by other authors. Nieswiadomy,[17] for example, stressed that hypotheses or research questions must be based on a propositional statement or statements from the theory. Polit and Hungler added that hypotheses are "predictions about the manner in which variables would be related, if the theory were correct and useful."[18(p92)]

Serlin[19] noted that scientists must state in advance what outcomes will be accepted as falsification of the theory under test, agreeing that no ad hoc appeals to unaccounted factors should be made. Looking toward the future, Serlin argued for "the appraisal of theories and the detection of progress if science is to take on an historical character, replacing the pitting of two theories against each other with a competition between research programs."[19(p369)]

The substantive use of structural equation modeling has been growing in psychology and other social sciences and is beginning to find some acceptance in nursing as well.

It should be noted here that the substantive use of structural equation modeling has been growing in psychology and other social sciences and is beginning to find some acceptance in nursing as well. One reason is that these confirmatory methods provide researchers with a comprehensive means for assessing and modifying theoretic models.[20]

In describing the findings of a study, Nieswiadomy[17] stressed the need to focus on explanations provided by the theory being tested. Further, she stated that any implications for nursing must be based on "the explanatory power of the theory."[17(p102)] The ability of a theory to explain study findings, or conversely, its inability to hold up as expected, is important in the determination of its value for clinical practice.

The preceding discussion has addressed some of the issues related to theory-testing research. From this review of the literature concerning theory-testing, 15 criteria were identified by the authors as essential for evaluating theory-testing studies. No attempt was made to include all criteria necessary for a comprehensive research critique. As Silva stated, just "because a study has met the evaluation criteria for theory testing does not necessarily guarantee that the study is sound. Other evaluation criteria, in addition to those related to theory testing, are important in assessing the overall quality of a study."[8(p4)] The criteria proposed by the authors and the rationale for the selection of each criterion follow.

CRITERIA FOR EVALUATING THEORY-TESTING RESEARCH

Criterion 1

The purpose of the study is to examine the empirical validity of the constructs, concepts, assumptions, or relationships from the identified theoretic frame of reference.

Rationale: It is important for the investigator to explicate the purpose of the study as being the testing of specific aspects of a theory or theoretic framework. Doing so assists the reader in determining that the investigator really did test a theory and did not just use it to organize the study or guide instrument selection. Readers need to know if the underlying tenets of the theory are valid, to avoid erroneous interpretations of data.

Criterion 2

The theoretic frame of reference must be explicitly described and summarized.

Rationale: Making explicit the tenets of a model or theoretic frame of reference accomplishes two purposes. First, readers can become familiar with the model and/or validate their perceptions of it. Second, readers can follow the logic (or illogic) of the author through the remaining steps of the report. Potential consumers must be able to determine a clear relationship between the model and the hypotheses to be tested.

Criterion 3

The constructs and concepts to be examined are theoretically defined.

Rationale: Theoretic definitions communicate the general meaning of the concept and permit the identification of empirical indicators.[14] The researcher must define the concepts or constructs on the basis of the theory so that accurate and valid measures will be obtainable.[17] Appropriate instruments are sometimes hard to find, and the researcher may elect to adapt a conceptual definition; however, this decision must then be explicated for the reader.

Criterion 4

An overview of previous studies that are based on the theoretic framework, or that clearly show the derivation of the concepts being tested, must be included in the review of the literature.

Rationale: The researcher must present evidence that he or she has conducted an extensive search of the literature for studies based on the theoretic frame of reference. The reporting of this research demonstrates that the investigator is thoroughly familiar with and understands previous theoretic work. Studies reported in the review of the literature should provide a logical base for the present study. In addition, Polit and Hungler[18] noted that the researcher should evaluate the theory before using it as the basis of a study.

Criterion 5

The research questions or hypotheses are logically derived from the definitions, assumptions, or propositions of the theoretic frame of reference.

Rationale: Research questions or hypotheses should be "explicitly formulated . . . to show, with the maximum exactness possible, relationships between the theory base for the research and the particular research study being conducted."[14(p156)] Theory-testing research permits the investigator to ascertain how accurately the theory is able to predict events in the empirical world. To determine the accuracy, the researcher poses hypotheses derived from the theoretic relationships that can be empirically tested and states them as questions or hypotheses.[21]

Criterion 6

The research questions or hypotheses are specific enough to put the theoretic frame of reference at risk for falsification.

Rationale: The goal of theory testing is to refute or falsify hypothetical relationships derived from the theoretical framework. "When a hypothesis states a relationship between empirical indicators, all assertions about that relationship must be testable."[22(p59)] "A falsifiable hypothesis must be sufficiently precise so that incompatible empirical results can be easily identified."[22(p60)] The findings must provide support for potential confirmation or disconfirmation of the theory.[18] Hypotheses must also be justifiable, that is, they must be derived from the theoretic frame of reference and their relationships supported by prior research findings.[18,23,24]

Criterion 7

The operational definitions are clearly derived from the theoretic frame of reference.

Rationale: To test a theory empirically, concepts must be clear and empirically observable.[22] If the purpose of the study is theory testing, then the operational definition of the concepts must be derived from the theory.[17] According to Jacobs,[12] one variant of testing theory is the selection of conceptual linkages from the parent framework for testing. The need for theoretically derived operational definitions is particularly important to this process.

Criterion 8

The design is congruent with the level of theory described in the theoretic frame of reference.

Rationale: To use the empirical approach, one must determine what sort of evidence will validly test the theory and must rule out competing hypotheses.[3] The evidence needed will be based on the extent to which

the theory is developed. If all phenomena have not been clearly described, then the theory-testing task would be to describe one or more phenomena. If the phenomena are described, but all relationships are not established, then the researcher's task would be to explore relationships among the phenomena. Once these relationships are known, theory testing would seek to validate them. Often, the theory-testing design is experimental, quasi-experimental, or correlational.[18,25]

Criterion 9

The instruments must be theoretically valid and reliable.

Rationale: To validate or refute theory, measurement of the theoretic concepts and their relationships must be made. Measurement via empirical instruments is fundamental to the testing of theoretic concepts and their relationships. Evidence of instrument validity and reliability for the concepts under study must be included in the research report. Such information permits the reader to determine whether the concepts of the theory were indeed measured correctly.[26]

Criterion 10

The theoretic frame of reference guided the sample selection.

Rationale: If one is conducting research to validate or falsify theory, selection of the sample must be theoretically guided. "Theory must determine in what ways the sample should be representative."[19(p366)] If the sample is not representative of the persons to whom the theory applies, the researcher will not have put the theory at risk for falsification.

Criterion 11

The statistics used are the most robust possible.

Rationale: The purpose of theory-testing research is to validate or falsify the theory; therefore, the strictest possible test must be used. Experimental, quasi-experimental, or exploratory designs are commonly associated with theory-testing research, and they necessitate the use of inferential statistics. Causal or structural equation modeling is the most stringent statistical test, as it allows the researcher to assess and modify the model.[3] However, the developmental level of the model will dictate the design and, accordingly, which statistic is the most robust for the study.

Criterion 12

The analysis of data must provide evidence for supporting, refuting, or modifying the theoretic framework.

Rationale: The empirical validity of a model must be determined so that one may use the theory upon which the study is based.[8] The analysis of data must also be congruent with the aims of the theory for appropriate interpretation of the data to occur.[14] Reynolds stated that "there are three possible outcomes for any research: either it supports the theoretical statement, it does not support the theoretical statement, or it is inconclusive."[15(p118)] The purpose of the analysis of data is an attempt to produce one of these outcomes. Data analysis, therefore, presents evidence to accept or reject the proposed hypotheses that allows a researcher to make conclusions regarding the validity of the model.

Criterion 13

The research report must include an interpretive analysis of the findings in relation to the theory being tested.

Rationale: For the reader to evaluate the theory in question, the researcher must provide an interpretive analysis of the data in relation to the explanations provided by the theory. Theory is only as strong as its tests of confirmability; therefore, the research report must present conclusions supporting, or failing to support, the theoretic relationships.[10,15]

Criterion 14

The significance of the theory for nursing is discussed in the report.

Rationale: Reynolds[15] stated that substantive significance is usually more important than statistical significance. The researcher must report the implications for nursing on the basis of the support provided by the explanatory power of the theory.[17] Without such information, the reader is unable to determine the potential usefulness of the theory for practice.

Criterion 15

Ideally, the researcher makes recommendations for further research on the basis of the theoretic findings.

Rationale: Although this is not really an essential criterion for theory-testing research, it is a vital step in the research process. First, the researcher should now be thoroughly familiar with the theory in question and therefore expert in the research needed to further its validity. Second, the researcher is most familiar with the data generated by the study and best able to ascertain directions that might prove fruitful for further investigation. Failure to report these recommendations could hinder future efforts to provide support or rejection of the theory, thereby retarding the development of nursing science. Therefore, the authors have elected to make this an essential criterion.

The research report must include an interpretive analysis of the findings in relation to the theory being tested.

APPLICATION OF EVALUATIVE CRITERIA

A number of studies were examined to evaluate the usefulness of the identified theory-testing criteria. Of these, six[27–32] were selected for critique, so as to demonstrate the application of the 15 criteria. These studies represented well-known nursing models and differed in design and statistical analyses. Fig 1 presents the studies examined during the process of determining the adequacy of the criteria.

The research reports selected for critique were representative of research explicitly testing theory. Each used a nursing theory or conceptual model to provide the framework from which concepts and propositions for testing are derived. Four of the reports[27–29,32] explicitly stated that they were testing concepts or propositions from a nursing theory or conceptual framework. Braden,[27] Kline-Leidy,[30] and Smith[32] demonstrated fulfillment of all 15 identified criteria. Each of these reports is an excellent example of theory-testing research. While Frey,[28] Frey and Denyes,[29] and Mishel and Braden[31] did

Study Author	1. Purpose	2. Theory	3. Construct/concept	4. Review of Literature	5. Question/Hypothesis	6. Specificity	7. Operational Definitions	8. Design	9. Instruments	10. Sample	11. Statistics	12. Analysis	13. Interpretation	14. Significance	15. Recommendations
Braden	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Frey	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–
Frey & Denyes	+	+	+	–	+	–	+	+	+	–	+	+	+	–	–
Leidy	+	+	+	+	+	+	+	+	+	–/+	+	+	+	+	+
Mishel & Braden	+	+	+	+	+	+	+	+	–/+	+	+	+	+	–	–
Smith	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Fig 1. Evaluation of the usefulness of the identified theory-testing criteria, using six selected studies. + = criterion met; – = criterion not met; –/+ = criterion partially met.

not fulfill each of the 15 criteria, they are to be commended for research aimed at expanding the knowledgebase for nursing.

RECOMMENDATIONS

First, researchers should identify theory-testing studies in their abstracts, publication titles, and library retrieval keywords. This strategy was initially proposed by Silva[8] to remedy the difficulty that investigators face in locating studies that have tested nursing models. By specifying the model or theorist in the title or abstract, the author will facilitate the dissemination of information about the validity and use of specific models.

Second, investigators should continue to study and refine theory-testing criteria. Through systematic effort to identify, apply, and evaluate criteria used in theory-based research, nurse scientists will accelerate development of the nursing database.

Third, nursing scholars should consider theory-testing criteria in relation to their individual research programs as well as to the research priorities for the nursing profession as a whole. Decisions about nursing versus interdisciplinary models, methodology, and level of theory development affect the quality of the research outcomes as well as the focal variables for study.

• • •

The authors have examined the concept of theory testing and some issues concerning theory testing in nursing. In addition, they

have identified from the literature 15 criteria that are proposed as essential for evaluating theory-testing research. To demonstrate the value of the criteria, the authors applied them systematically to six studies in which concepts deducted from nursing models were tested.

This article builds on the work of Silva[8] and others[14,16-20] who have explicitly delineated formative evaluation criteria. The authors put forth this effort in the spirit of challenge and hope that nurse scientists will continue to improve the quality of theory-testing research.

REFERENCES

1. Engstrom JL. Problems in the development, use and testing of nursing theory. *J Nurs Educ.* 1986;23:245–251.
2. Schlotfeldt RM. Resolution of issues: An imperative for creating nursing's future. *J Prof Nurs.* 1987:136–142.
3. Hanley C. A problem of theory testing. *Int Rev Psycho-Anal.* 1983;10:393–405.
4. *ANA Social Policy Statement.* Kansas City, Mo: American Nurses Association, 1980.
5. Moody LE. *Advancing Nursing Science Through Research.* Newbury Park, Calif: Sage, 1990.
6. Chinn PL. From the editor. *ANS.* 1984;6:ix.
7. Moody LE, Wilson ME, Smyth K, Schwartz R, Tittle M, Van Cott ML. Analysis of a decade of nursing practice research: 1977–1986. *Nurs Res.* 1988;37(6):374–379.
8. Silva MC. Research testing nursing theory: State of the art. *ANS.* 1986;9(10):1–11.
9. Thomas BS. *Nursing Research. An Experiential Approach.* St. Louis, Mo: Mosby, 1990.
10. Treece EW, Treece JW. *Elements of Research in Nursing.* St. Louis, Mo: Mosby, 1986.
11. Seaman CHC. *Research Methods. Principles, Practice and Theory for Nursing.* Norwalk, Conn: Appleton & Lange, 1987.
12. Jacobs MK. Can nursing theory be tested? In: Chinn PL, ed. *Nursing Research Methodology.* Rockville, Md: Aspen Publishers, 1986.
13. Fawcett J. Conceptual models and theory development. *J Gynecol Neonat Nurs.* 1988;17(6):400–403.
14. Chinn PL, Jacobs MK. *Theory and Nursing. A Systematic Approach.* St. Louis, Mo: Mosby, 1987.
15. Reynolds PD. *A Primer in Theory Construction.* New York, NY: Macmillan, 1971.
16. Walker LO, Avant KC. *Strategies For Theory Construction in Nursing.* 2nd ed. Norwalk, Conn: Appleton & Lange, 1988.
17. Nieswiadomy RM. *Foundations Of Nursing Research.* Norwalk, Conn: Appleton & Lange, 1987.
18. Polit DF, Hungler BP. *Nursing Research. Principles and Methods.* Philadelphia, Pa: Lippincott, 1987.
19. Serlin RC. Hypothesis testing: Theory building, and the philosophy of science. *J Counseling Psychol.* 1987;34(4):365–371.
20. Anderson JC, Gerbing DW. Structural equation modeling in practice: A review and recommended two-step approach. *Psychol Bull.* 1988;103(3):411–423.
21. Woods NF, Catanzaro M. *Nursing Research. Theory and Practice.* St. Louis, Mo: Mosby, 1988.
22. Fawcett J, Downs FS. *The Relationship of Theory and Research.* Norwalk, Conn: Appleton-Century-Crofts, 1986.
23. Phillips LRF. *A Clinician's Guide to the Critique and Utilization of Nursing Research.* Norwalk, Conn: Appleton-Century-Crofts, 1986.
24. Wilson HS. *Research in Nursing.* Menlo Park, Calif: Addison-Wesley, 1985.
25. Campbell DT, Stanley JC. *Experimental and Quasi-experimental Designs for Research.* Boston, Mass: Houghton Mifflin, 1963.
26. Rew L, Stuppy D, Becker H. Construct validity in instrument development: A vital link between nursing practice, research, and theory. *ANS.* 1988;10(4):10–22.
27. Braden C. A test of the self-help model: Learned response to chronic illness experience. *Nurs Res.* 1990;39(1):42–47.
28. Frey MA. Social support and health: A theoretical formulation derived from King's conceptual framework. *Nurs Sci Q.* 1989;2(3):138–148.
29. Frey MF, Denyes MJ. Health and illness self-care in adolescents with IDDM: A test of Orem's theory. *ANS.* 1989;12(1):67–75.
30. Kline-Leidy N. A structural model of stress, psychosocial resources, and symptomatic experience in chronic physical illness. *Nurs Res.* 1990;39(4):230–236.
31. Mishel M, Braden C. Finding meaning: Antecedents of uncertainty in illness. *Nurs Res.* 1988;37(2):98–103, 127.
32. Smith M. Testing propositions derived from Rogers' conceptual system. *Nurs Sci Q.* 1988;1(2):60–67.