

COMMENTARY

No Research Versus Research With Compromised Results: A Study of Validation Therapy

SUSANNE S. ROBB • CHARLES E. STEGMAN • MARY OPAL WOLANIN

Experimental designs have become the cornerstone of experimental research in the health sciences. Control of extraneous variables and threats to internal and external validity are well-known attributes of this research (Cook & Campbell, 1976). Unfortunately, in many instances such methodological rigor cannot be achieved, and serious design problems follow (Flay & Best, 1982). One of these instances may be found in institutional care settings serving elderly resident populations.

One extreme solution to these problems is to not conduct research. Perhaps this solution is one reason that gerontological research is infrequently performed in institutions. However, were everyone to adopt this strategy, nurses would practice in a research-based vacuum.

Agencies engaged in long-term care of the elderly tend to adopt programs without valid information about their effectiveness. Resources required to implement them, in the context of controlled trials, are rarely available. The programs begin, run for a while, and then vanish. Only the developers who market them are clear beneficiaries! Geriatricians and gerontologists are years away from curing problems that the programs attempt to alleviate. Closer-to-ideal

research environments in long-term care are even further away. It is debatable whether evaluations, done under ideal conditions, will produce programs that can be readily adopted in suboptimal settings.

This article reports a study of validation therapy, a popular program that is threatening to surface in many long-term care facilities in the absence of research-based information about its effectiveness. By the time this study was completed, it was so flawed that the results were severely compromised. The real value in publishing this report rests in its potential to encourage others to evaluate validation therapy, using flawed designs if necessary, so that a body of valid information can emerge, based on conceptual generality.

Replication to achieve conceptual generality has been a fundamental tenet of science. Replication is not as simplistic as repeating an experiment the same way. The science of physics is not advanced by having students perform the same classical experiments over and over again; these students are learning about experimentation, not increasing core scientific knowledge. Replication entails the ability of ideas or results to hold up under generalized conditions. Hendrick and Jones (1972) argued that the basic premise of conceptual generality is that when the same results are found with different subjects, situational factors, and other variations in study procedures, the aggregate results are accepted with greater confidence than the results from any single study. The idea of conceptual generality represents a potential solution to the problem of obtaining valid results from studies that lack methodological rigor because they focus on human behavior in complex health care environments.

Research Background: Reality orientation (Taulbee & Folsom, 1966) achieved acceptance as a therapy for the severely disoriented elderly before a number of important questions were answered, such as whether the process or content was responsible for the effects produced. One major objection to reality orientation concerns the degree to which emphasis on orienting residents in long-term care settings to time, place, and person reflects the realities and demands of the institutional environment (Bernstein & Dvorkin, 1978; Carroll & Gray, 1981). For example, is it functionally adaptive to be continually aware of time within an institutional setting characterized by the absence of activities to fill time? Johnson (1980), Mitchel (1980), and Zepelin,

SUSANNE S. ROBB, PhD, RNC, FAAN, is associate chief, Nursing Service for Research, VA Medical Center, University Drive, Pittsburgh, PA.

CHARLES E. STEGMAN PhD, is associate professor of educational research, University of Pittsburgh, Pittsburgh, PA.

MARY OPAL WOLANIN, MPA, RN, is associate professor emeritus, University of Arizona, and consultant in geriatric and gerontological nursing, Tucson, AZ.

Accepted for publication August 22, 1985.

The assistance of Ronald H. Rabold, MSW, ACSW, social work supervisor, Cynthia J. Gifford, BSN, RN, and Sigrid V. Epstein, BSN, RN, staff nurses, and Mary Winston, BSW, social work associate, who conducted the therapy, is acknowledged.

NURSING RESEARCH

Initially, 16 subjects were random-
ly assigned to the control group and
20 to the experimental group. Ran-
domization does not guarantee small
sample sizes. Fortunately, the two
groups in this study had similar mean
ages (80 and 81 years), and 12 and 13
members, respectively, had diagnoses
indicative of dementia. The experi-
mental group had a slightly less
favorable mean mental status score
(11 vs. 13 points) and a longer mean
length of stay (3.8 vs. 2.8 years).

Attrition: Subject loss, due to ill-
ness and death, is to be expected in all
elderly populations. Researchers often
treat oversample to compensate for this
type of attrition. But, when all eligible
subjects are included in a study, the
final sample size is impossible to predict
because there is no control over the
number of subjects who drop out.

Control Conditions: Ideally, sub-groups are treated exactly the same except for the program that is being evaluated. In this study, the treatment-giving component than other conditions of programs such as media-ions, where double-blind and placebo-procedures can be used.

This procedure was approved by the human rights committee only because of the minimal risk inherent in the study procedures and the difficulty experienced in locating relatives or guardians. The committee realized that serious legal and ethical questions existed concerning the use of "assent" in place of actual written consent. When the revised code of Federal Regulations, concurred with the protection of human subjects (National Research Act, 1968), was adopted in 1983, the code did not include a section dealing with research on the elderly and/or mentally disabled in effect, did not mention the procedure.

Willfulness to participate was often
determined by absent, rather than written
consent. In this study, absent was
defined as the absence of verbal or
nonverbal objection to possible par-
ticipation in a program involving
group discussions of past experiences.

The study group included all residents who were (a) 60 years and older, (b) moderately to severely disabled, (c) suffered neurological disorders in- tellec- tual, (d) Alzheimer's disease, Pick's disease, Huntington's chorea, or a cerebrovascular accident within the past 6 months, and (e) unlikely to be discharged during the next 12 months. Head nurses and physicians were asked independently to nomi- nate all residents who met the study criteria. Of the 398 residents, 60 were included; 36 met the eligibility cri- teria.

An interrupted time-series design with switching dependent variables represents one solution to the problem of inadequate numbers of subjects. Although there are several sources of invalidity within this design (Cook & Campbell, 1976), a major problem concerns the repeated measurement. This burden places a strain on staff resources imposed by the available resources for the validation therapy study were not sufficient to support any more than single-pre- and postmeasurements. Therefore, in the current study, a pretest-posttest, two-group (control and experimental) design was initially planned with ran-

Sample. Questions about adequate sample size and selection procedures are often difficult. In long-term care settings, the number of residents who meet the study criteria is usually smaller than the researcher would prefer to obtain. In this situation, high statistical power is usually attained by randomly selecting an appropriate number of subjects from a large population.

Wolle, and Kleinplatz (1981) all questioned the benefit of severely disorientated for elderly disorientation after she observed that set aside persons. Feil (1967) proposed an alternative therapy for elderly clients to respond to orientation that set aside persons. Her treatment, generally referred to as validation therapy, is based on the premise that there is logic behind the behavior of this therapy; the goal is to help the individual relate behaviors to logic behind them. Awareness of reality is not the goal of this therapy; the goal is to help the individual reach the inner reality in order to reach the individual (Wetzler & Feil, 1978). Specific procedures and techniques are described by Feil in a number of sources (E. Feil, 1972, 1978; N. Feil, 1982a, 1982b). This therapy had been evaluated only by its developers using mostly female subjects. Their evaluation methods have not been published, however, and their results are therefore based on clinical experience.

trol group (loss of 4) and 15 subjects (loss of 5) in the experimental group.

The program was offered twice a week for the 9 months from September through May. Six subjects who were originally assigned to the experimental group did not attend most of the scheduled sessions because of episodic acute illnesses ($n = 5$) or extremely disruptive behavior during group meetings ($n = 1$). This partial treatment group was not the same as the control or treatment groups. Some researchers opt to drop such a group entirely. Others include these subjects with the control group or, worse yet, count them as treatment subjects. The solution favored here was to include them as a special subgroup. Thus, data analysis included 12 persons in the control group, 9 in the experimental group, and 6 who had been randomly assigned to the experimental group, but who failed to attend a majority of the meetings. Compared to persons who remained in the experimental or the control group, those who dropped out were older, had longer lengths of stay in the facility, and demonstrated lower mental status scores.

Measures: The selection and/or development of appropriate measurement procedures are vitally important to behavioral science research. And yet, in many institutional settings there is little support for this type of research. In these cases, researchers are often forced to use ill-defined subjective measures of program effectiveness from existing measures that may or may not be adaptable to the current situation. Scarce personnel resources, lack of suitable instruments, and finite subject tolerance for testing also limit the number of dependent variables that can be considered in any study. Theory and previous research dictate the important variables to consider.

Of all the treatment effects reported by Feil and Wetzler (1979) and Wetzler and Feil (1978), three were selected for study. Mental status, morale, and social behavior were chosen because of the researchers' beliefs that change might occur more rapidly in these areas. Also, these variables could be quickly measured with few disruptions to the residents and the nursing staff. The developers and sole evaluators of validation ther-

apy had not used procedures with established reliability, validity, and sensitivity. Therefore, tools with at least partly established reliability and validity were selected to measure each variable. Mental status of the subjects was measured by the Mental Status Questionnaire (MSQ) developed by Fishback (1977). Morale was measured with the Philadelphia Geriatric Center Morale Scale (PGCMS) developed by Lawton (1975). The Minimal Social Behavior Scale (MSBS) by Farina, Arenberg, and Guskin (1957) was used to evaluate social behavior.

The MSQ (Fishback, 1977) is a 36-item questionnaire that combined the Kahn, Goldfarb, Pollack, and Peck (1960) and Pfeiffer (1975) mental status questionnaires to give greater latitude and accuracy in determining degrees of senile dementia. Some degree of validity had been established in that "scores correlate highly with clinical findings" (actual coefficient not reported) (Fishback, 1977, p. 167). Reliability was not established for the combined instrument, but

Pfeiffer's Short Portable Mental Status Questionnaire (SPMSQ) had test-retest correlations of .82 and .83 for two groups tested at 4-week intervals. In terms of validity, the SPMSQ was significantly, $p < .001$, correlated with clinical diagnosis of organic brain syndrome in two study groups. The MSQ was found to be "highly related to psychiatrists' clinical evaluations of the presence and degree of chronic brain syndrome" (actual coefficient not reported) (Kahn et al., 1960, p. 328).

The PGCMS is a 17-item, 3-factor scale (tranquility, attitude-toward-own-aging, and lonely-dissatisfaction) that yields an overall measure of morale. Reliability statistics were as follows: KR 21 = .81 (matched items); $r = .75$ test-retest (5 weeks). Validity studies indicated a moderate correlation with interviewers' life satisfaction ratings ($4 = .57$) (Lawton, 1972, 1975), as well as high correlations with other morale scales: LSIA = .74; Dean = .57; Kutner = .73; LSIZ = .79; and LISA-A = .77 (Lohman, 1977, p. 74).

FACULTY OF NURSING THE UNIVERSITY OF ALBERTA

The University of Alberta offers challenging opportunities to participate in expansion of Basic (BScN) Program, Master of Nursing Program and nursing research endeavours and in developing a doctoral program in nursing. Joint appointments with health care agencies available.

Six positions available may be at Associate or Assistant Professor rank, dependent on qualifications. Doctoral degree, preferably in Nursing; relevant clinical, teaching, and research experience; and current registration/licensure as RN required. Master's degree in Nursing preferred. Clinical background in Community Health Nursing; Gerontological Nursing; Maternity Nursing; Medical-Surgical Nursing; and Nursing of Children.

Salary ranges: Associate Professor—\$38,170-\$55,450
Assistant Professor—\$30,316-\$43,780

The University of Alberta is an equal opportunity employer

Submit curriculum vitae and names of three references to:

Dr. Jannetta MacPhail, Dean
Faculty of Nursing
The University of Alberta
3-121 Clinical Sciences Building
EDMONTON, Alberta, Canada, T6G 2G3.

Table 1. Differences in mean scores and standard deviations for three subscale groups between pre- and posttest on Three Measures

The empathatic model used in validation therapy placed unusual demands on the members/leaders who needed to be constantly active—physically and mentally—during sessions. On occasions, leaders found themselves unable to elicit responses from the men or involve them in the task at hand. Frustrated leaders relied from the men or involve them in the task at hand. Frustrated leaders relied heavily on each other for support. A second source of frustration occurred when other personnel impeded the therapy team was making a "futile" effort to assist a "helpless" group. Costs could possibly be reduced by shifting the program to clinical units, particularly those with large proportions of the kinds of residents believed by Feil and Wetzler (1979) to benefit from the "old-old," and without additional staff would be required. It validation therapy were conducted by all personnel as an integral part of clients' treatment plans, no cover the extensive and ongoing costs would then shift to participants in the education program to all personnel.

did not have sufficient time to support volunteer escorts. For the program to continue as a discrete entity, additional personnel would be needed. However, a full-time validation therapist might be imitated in ability to provide effective programming. The nature of the therapy is such that one person might become extremely bored with conducting more than two or three sessions a day for 4 or 5 days each week.

procedures under a randomization model is more justifiable (Lehmann, 1975). There was little difference between the pretest-posttest results or the steepest results between the groups on the three measures (Table I). The steepest differences at the .05 level between all groups revealed no significant differences at the .05 level between all groups. However, using the Kruskal-Wallis one-way ANOVA on the post-treatment scores revealed no significant differences between all three variables.

care abilities (such as dressing, toileting, or bathing). However, many of the residents did demonstrate changes in their social behaviors. Increased demands on the nursing staff for cigarettes, conversation, and so forth prompted at least one nurse to comment, "If these men keep returning so riled up, we'll just have to increase their tranquilizers." In many instances, staff retained negative attitudes despite having received materials describing the therapy and its potential benefits. Part of this problem stemmed from differences in goals held by the therapy team versus goals held by the staff. The latter group expected bowel and bladder continence to be restored, but the therapy team was satisfied with subjects' gains in social behavior or recognition of feelings.

Implications: Before results of this study can be used for decision-making, they should be integrated with those from other evaluations. Given the limitations imposed on doing research in long-term care facilities, it

is doubtful if any single study will ever sufficiently control the threats to internal and external validity to warrant making final decisions about the worth of a new treatment program.

A question that researchers must face as they design their studies is, "At what point does inability to control threats to internal and external validity render the research worthless?" In adapting true experimental designs to fit the situational realities of long-term facilities, a number of compromises are often made. Each of these must be acknowledged in reports and the results discussed in light of these limitations. Although the researcher may not be justified in making causal inferences based on the study, the results can still inform decision-makers about some aspects of the phenomenon studied and point the way to better studies. The accumulation of results through techniques like meta-analysis (Glass, McGraw, & Smith, 1981; Slavin, 1983) may also lend strength to the conceptual generality of the results. Results are reported here in an effort to facilitate such comparisons. ■■■

MOVING?

GIVE US
YOUR NEW ADDRESS
SIX WEEKS AHEAD

To ensure prompt service when you change your address or whenever you write to us about your NURSING RESEARCH subscription, please include your address label from a recent issue of NURSING RESEARCH. Address your correspondence to Subscription Department, NURSING RESEARCH, 555 West 57th St., New York, N.Y. 10019-2961.

Affix address label here when sending address change. Clearly print NEW address. The expiration date of your subscription is toward the right end of the top line of your address label, followed by month, e.g. 85 Sept.

Name

Address

City

State Zip

**springer
publishing co**

**Annual Review of
Nursing Research, Vol. 4**

**Harriet H. Werley,
Joyce J. Fitzpatrick,
Roma Lee Taunton, Editors**

NEW

This volume draws together and critically reviews research in four categories: nursing practice, nursing care delivery systems, nursing education, and the profession of nursing.

Partial contents: • *Maternal Anxiety in Pregnancy* • *Preschool Children* • *Menopause* • *Aging: Gerontological Nursing Research* • *Bereavement* • *Evaluation of Primary Nursing* • *Nurse-Midwifery Care: 1925-1984* • *Faculty Productivity* • *Nontraditional Nursing Education* • *Computer-Aided Instruction in Nursing Education* • *Nursing's Heritage* • *Nursing Education Research in Canada, 1986* / Pre-pub \$31.00

**Essentials of Nursing
Research/3rd Ed.**
Lucille Notter

Widely adopted in nursing schools across the country, this book is described by Barbara Stevens in her foreword as: "A classic in nursing research.... It makes the complicated process of nursing research clear, comprehensible, and even exciting for the uninitiated." 192pp / 1983 / soft \$11.50

**First-Time Motherhood:
Experiences from Teens to Forties**
Ramona Mercer **NEW**

Presents unique research into the ways women experience first-time motherhood, with revealing vignettes of selected cases. 1986

**The Clinical Career Ladder:
Planning and Implementation**
**Laura Merker, Kathleen Mariak, and
Dana Dwinells**

A complete guide to implementing a career ladder program for clinical nurses in hospital and other health-care facilities. Describes the benefits in terms of staff retention, professional development, and improved resource management. 176pp / 1985 / \$15.95 soft

Please send _____

Enclosed is _____ (Add postage: \$1.60 first book, 40¢ thereafter. In NY add tax) N14

Name _____

Address _____

City/State/Zip _____

Springer Publishing Co.
536 Broadway, New York, NY 10012

NURSING RESEARCH

CALENDAR OF RESEARCH EVENTS

- PLATT, B. H., & BEEST, J. A. (1982). Overcoming design problems in evaluating health behavior problems. *Evaluation and Professions*, 5(1), 43-69.

FOX, R. N., & VENUTRA, M. R. (1983). Small-scale administration of instruments and procedures. *Nursing Research*, 32(1), 122-125.

LAWTON, M. P. (1975). The Philadelphia Geriatric Center Morale Scale: A revision. *Journal of Gerontology*, 30(1), 55-69.

GILLES, G., V. MCNAUL, B., & SMITH, M. L. (1981). Metronidazole in social research. *Health Psychology*, 1, 197-212.

LAWTON, M. P. (1977). Correlations of life satisfaction, morale and adjustment measures. *Journal of Gerontology*, 32(1), 73-75.

LOHMAN, N. (1977). Comparison of the salt-sugar ratio of meat-analogs in educational research. *Journal of Gerontology*, 32(1), 76-81.

SILAVAN, R. E. (1983, April). Metronidazole: use of meat-analogs in educational research. Paper presented at the Annual Convention of the American Geriatrics Society, San Diego, CA.

SLAVIAN, R. E. (1983, April). Metronidazole: use of meat-analogs in educational research. Paper presented at the American Geriatrics Society, San Diego, CA.

ROBB, S. S. (1983, July). Third annual survey of major client satisfaction and problems of older adults and potential donors. *Chiefs, Nursing Service for Reebach, VA Medical Center, Nursing Service for Reebach, VA* (Available from Superintendent, St. Bobb, Associate Director, C. Pittsburg, PA 15240).

SLAVIAN, R. E. (1983, April). Metronidazole: use of meat-analogs in educational research. Paper presented at the American Geriatrics Society, San Diego, CA.

LOHMAN, N. (1980, November). Reality orientation in the nursing home: A test of effectiveness of reality orientation for the elderly aged. *Long Term Care and Health Services Administration Quarterly*, 4(2), 103-114.

WETZLER, M. A., & FEIN, N. (1978). Manual for Implementing the Bell Method. Cleveland, OH: Edward Bell Productions.

ZEPPELIN, H., WOLFE, C. S., & KLEINERMAN, F. (1981). Evaluation of a year-long reentry orientation program. *Journal of Gerontology*, 36, 70-77.

PEFFER, E. (1974, 3 46) (March 8, 1983). National Research Act. *PUBLIC LAW 98-348, 92 Statute of the American Geriatrics Society, San Diego, CA.*

KRANZ, R. L., GOLDSTEIN, A. I., POLLACK, M., & LOGICAL SOCIETY, SAN DIEGO, CA.

PECK, A. (1986). Brief objective measures for the determination of mental status in the elderly. *American Journal of Psychiatry*, 143, 328-329.

PECK, A. (1986). Brief objective measures for the assessment of elderly patients in psychiatric institutions. *Journal of the American Geriatrics Society*, 34, 1177.