DOCUMENT RESUME

ED 244 984 TM 840 306

AUTHOR	Herman, Joan L.
TITLE	Serving the Mutual Needs of Research and Practice: The Methodology Project Example. Draft.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Apr 84
NOTE	12p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Achievement Gains; *Educational Practices; Elementary Secondary Education; Evaluation Methods; Researchers; *Research Methodology; *Research Utilization; Scores; State Programs; Testing; Testing Programs
IDENTIFIERS	Center for the Study of Evaluation CA; Curriculum Test Overlap

ABSTRACT

This paper presents two examples of how methodological research in testing and evaluation simultaneously served the needs of both researchers and practitioners. The first example focuses on Center for the Study of Evaluation (CSE) research in writing assessment. This research investigated issues in domain definition, cognitive processing requirements, rater and scale stability, and response mode. The analytic scales resulting from this research were implemented, with CSE assistance, in district and state assessment programs. These implementations provided data to further the research. The second example focuses on theoretical research in integrating testing and instruction. A school district's desire to raise students' test scores provided the opportunity to reality test some of the research ideas. The result was higher test scores for the district's schools, and important data for the researchers. (Author)

ED244984

DRAFT

Serving the Mutual Needs of Research and Practice: The Methodology Project Example

Joan L. Herman

Center for the Study of Evaluation
UCLA

The research presented herein was performed pursuant to a grant from the National Institute of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

Paper Presented at the Annual Meeting
of the American Educational Research Association
New Orleans, LA
April 1984

Serving the Mutual Needs of Research and Practice:

The Methodology Project Example

Joan L. Herman
Center for the Study of Evaluation
UCLA

This paper presents two examples of how methodological research in testing and evaluation simultaneously served the needs of both researchers and practitioners. The first example focuses on CSE research in writing assessment. This research investigated issues in domain definition, cognitive processing requirements, rater and scale stability, and response mode. The analytic scales resulting from this research were implemented, with CSE assistance, in district and state assessment programs. These implementations, in turn, provided data to further the research.

The second example focuses on theoretical research in integrating testing and instruction. A school district's desire to raise students test scores provided the opportunity to reality-test some of the research ideas. The result was higher test scores for the district's schools, and important data for the researchers.

The gap between research and practice if an oft cited abyss which has been the subject of longstanding conjecture and concern. Researchers in various disciplines have expended considerable energies in exploring the reasons for the gap; they have discovered many. They have dissected the values, norms, expectations, roles, functions, and incentive structure within the research and practice worlds and have found disparities. They have observed that the problems attacked in many research studies often appear insignificant to human service practitioners (Rosenblatt, 1968) while attempts to use research findings in real world settings and to develop realistic approaches to practitioner problems are berated, viewed as crass and inelegant by "pure" scientists (Archibald, 1968; NSF, 1969). The latter view seems to derive from university sanctions which function to discourage the development of applied knowledge and/or the real world applications of theoretical knowledge, sanctions which encourage the development of knowledge for its own sake and the publication of knowledge as a sole end goal (Glaser et al, 1983). Some have drawn attention to the unfortunate end product of this "two world" syndrome: the general tendency of each side to stereotype, belittle, and reject the work of the other and to avoid and/or be overtly hostile toward each other (Schmuck, 1968).

While there is little doubt that the two worlds of research and practice do exist to some extent as separate entities with different needs and goals, there is also evidence that the breadth of the chasm is exaggerated, representing stereotypes more than reality. Weiss and Weiss (1980), for example, examined the attitudes of social scientists and decision-makers toward one another and toward the roles of social science research in public policy research in the field of mental health. They found that both groups valued research and agreed on what makes it useful; they did, however, disagree on the importance of specific characteristics of research. Interestingly, in rating their projections of policymaker's research criteria, social scientists' ascriptions were more practical and less methodologically sound than those to which policy-makers actually subscribed. In other words, researchers underestimated the policy makers research savvy and their attitudes toward methodological rigor. Such lingering stereotypes and the condescension they often represent do not facilitate the use of research in practice nor do they stimulate the kinds of joint efforts that are needed to solve pressing educational problems. The examples I'd like to present today illustrate the merits of moving away from the view that we straddle opposite ends, or at least fairly distant points, of a continuum and explore complementary needs and mutual problems.

The mutual needs perspective utilizes many of the central tenets of basic models of dissemination/diffusion. Like the RD&D model, which Marv has discussed, a concrete and well developed research idea is central, but the outcomes are multiple, the road is not always directly linear, and the audience, rather than passive, is active and collaborative. Like the problemsolving model, the perspective offered here also starts with the user's/practitioner's needs as a beginning point for utilization, with diagnosis as an essential first step in the search for problem solutions. But contrary to this model's non-directive stance, it asserts that users/ practitioners can benefit directly from the experience of research and ought to draw on external resources as well as internal problem solving processes for direction. Similar to the action research model that Adrianne has discussed, the mutual needs perspective presented here concurs on the importance of collaborative research within and by an organization. The research and its methodology are influenced by the on-going activities of the organization with which they co-occur and the research is a continuous process of research, action, evaluation, and more research enlightened by the evaluation findings. However, unlike this model, the outcomes of the research meet multiple needs and are directly linked to research aims beyond the organization and beyond the user. Finally, like the social interaction model, we of course acknowledge and try to be sensitive to the complex and intricate set of human relationships, structures and processes that are involved in knowledge utilization, technology transfer and change.

The mutual needs perspective sees research and practice as two overlapping systems; the overlap reflects interaction on an R&D problem of mutual concern. The practice system presents a need which is defined and refined through discussion and interactions; the research system brings a research-based solution and a planned variation, also based on discussions and interaction and tailored to the defined need. The solution ideally produces better practice and likewise provides findings which further the aims of research. In other words, research and practice come together to solve a problem of common concern and both leave with tangible benefits which are appropriate to their relative action worlds. (See Figure 1).

Figure 1

Let me provide a couple of examples from CSE's research program in testing and evaluation methodology. First, with funding from the National Institute of Education, CSE conducted extensive research on issues of constructed response assessment in general and writing assessment in particular. The research examined critical cognitive variables in defining the assessment domain and examined the effects on student performance of different genres, response modes, and prompt modalities; the research also investigated reliability and validity issues related to various scoring rubrics, training and scoring procedures. (See Quellmalz et al, 198 ; Quellmalz, 197 ; Spooner-Smith, 197 ; Pitts, 197 ; Smith, 197). As outcomes of this multi-year research effort (in addition to the traditional research reports and journal publications) analytic rating scales were developed for assessing narrative and expository essays, as were writing prompts in verbal and pictoral modalities and sophisticated training and rating procedures.

With the spread of minimum competency testing and increasing concerns about students' basic writing proficiencies, many districts and states needed advice and procedures to conduct assessments in this area: CSE was ready and willing to provide such advice and participated in the adaptation of CSE scales for various assessment programs, efforts which were funded by the LEA's and SEA's. From the standpoints of those served, CSE was providing primarily a technical assistance function. We worked with LEA's and SEA's to adapt our scales and procedures to meet their needs and constraints, trained their teachers and personnel to implement the adapted system and left a local capacity for continuing implementation. However, the technical assistance served another function for CSE purposes. Planned variations were embedded in nearly every effort, enabling replications and extensions of the NIE funded research. Within these variations, for example, the effects of picture vs. verbal prompting strategies were re-examined with various types of students; the phenomenon of rater drift and its likely control were further explored; cost-benefit analyses of single versus double ratings were conducted; and strategies for increasing the cost feasibility of analytic scoring systems were documented. These real world try-outs also allowed CSE to demonstrate the value of domain-referenced test methodology, its utility for meeting the evaluation needs at various administrative levels, and its promise for improving educational practice. By using an area of common concern, writing assessment, researchers and practitioners collaborated and both groups emerged wth what they wanted and needed. (Writing performance, in fact, improved in some districts after the installation of the CSE writing assessment program.)

As a second example, let me describe how our theoretical work in criterion and domain referenced testing and our conceptual work in integrating testing and instruction enabled a school district to solve a pressing political problem and permitted us to reality test some of our research ideas. A local school district approached us for help with a relatively common problem: their district, in fact every school save one, consistently scored below expectation on the state assessment, their school board was growing increasngly unhappy with this test performance, and district administration was under tremendous pressure to raise students' scores. We met with district administrators, analyzed the test scores on a district level, indicated areas of general strength and weakness, and suggested skills which held the greatest promise for improving the district's overall performance. The district found this exercise useful and wanted it extended on a school-by-school basis: they were making each principal responsible for his/her school's improved performance and wanted us to meet with each principal to analyze instructional strengths and weaknesses and make recommendations for action. Supplementing this school-by-school approach, and reflecting CSE's research orientation, we suggested that the district examine the match between the district curriculum and the state assessment and develop instruction and practice exercises where gaps existed.

The match we found between the state assessment and the district's instructional program and the results of our interventions are worth a short digression because they vividly demonstrate the need for and the potential impact of research in practice. First, we carefully compared the district's instructional program in reading with the test specifications. This involved a thorough analysis of the teacher's manual and accompanying student workbooks at two assessed grade levels for the basal series being used in the district, including a tally of the opportunities provided for direct instruction, appropriate practice, and other less directly related practice. The results were quite startling. The series included appropriate practice for less than half the objectives included in the test and virtually no guidance for providing students with direct instruction in comprehension; workbook exercises were similarly lacking instruction, results which are disturbing for a number of reasons. It is axiomatic that students will have a great deal of difficulty performing well on content which they have not been taught (or even assuming some of the test content might be taught in previous grade level texts, content to which they have been recently exposed). Further, given extensive research bases attesting to the importance of direct instruction and indicating effective strategies for teaching comprehension, it is frustrating that the texts provided so little guidance in direct instruction, another vivid reminder of the gap between research and practice.

As a result of this analysis, we developed practice exercises for all objectives covered by the test. We also analyzed the previous year's test results on a school-by-school basis and made individual recommendations for improving students' test performance. These recommendations suggested where the school might best put its effort in order to get the highest return, including such things as teaching to areas of relative weakness, teaching skills which were more amenable to instruction, teaching skills which were assessed across a couple of subject areas (e.g., suffixes were assessed in both reading and in language) and, less obviously, teaching skills where a small increment in raw scores would produce a considerable gain in standard score. We also made some more general suggestions about assuring that students and teachers knew that the test was important and were geared up for it, trying to make sure that students had a good night's sleep the night before test day, and testing proctoring procedures to make sure students were in fact attending to the test (and not, for example, gazing into space and/or making random designs on their answer sheets). These recommendations were delivered in individual meetings with each principal and were reiterated in follow-up letters. The result? While we hesitate to attribute cause, the district's scores went up on the next assessment and in fact all schools except one scored at least within the expectation band, a success story which clearly had benefits for the school district. CSE derived sizeable benefits as well, not the least of which was a good story to demonstrate to funders, legislators, and others the potential impact of research on practice and the efficacy of CSE's research in particular.

What can we learn from these two success stories about how to facilitate the impact of research on practice? The examples illustrate some of the principles cited in the literature for facilitating knowledge utilization and the impact of research on practice. (See, for example, Glaser, 1983; Hodgkin et al, 1973; Zaltman and Deshpande, 1979; Stolz, 1978.) Of critical importance is the identification of the problem/research area to be addressed: it must reflect the interest and concerns of those affected, including both the researchers and the practitioners, with an emphasis on the latter. In fact, one of the likely key ingredients is that practitioners have a pressing management problem that can be addressed by specific research interests.

Following on the importance of the practitioner's concerns, it also appears important that the project is initially requested or proposed by the practitioner rather than by the researcher, demonstrating a proactive stance that was evident in the two examples cited above. Practitioners' needs are preeminent, and in a related principle, it is crucial that research-based solutions are tailored to those needs. In the writing activities the CSE writing scales, in each instance, were modified to reflect local concerns and instructional priorities, while in the testing and instruction example the solution was customized for the district's constraints, desires, and administrative style. Flexibility and local adaptation, in other words, are all important attributes that help bridge the gap between research and practice, attributes which have been well documented in the literature (Berman and McLaughlin, 197).

Local needs and adaptation options are explored and negotiated during extensive interaction and collaboration between the researchers and the practitioners. Such personal interaction is necessary not only to assure that interventions are sensitive to local needs but also to create the trust and open communication that insures that solutions are perceived as credible and relevant to the realities of the practical situation.

The power of such collaborations should not be underestimated. They are visible evidence of both researchers' commitment to practice and their effectiveness in improving practice, demonstrations which not only increase the credibility of educational research but also contributed to its resource base and political support. By bridging the gap between research and practice, one uses scarce resources to their fullest (why reinvent the wheel in practice), creates networks which in turn generate new resources, and builds productive collaborative relationships which contribute better and real solutions to important educational problems.