
Assessing Program Outcomes: Rationale and Benefits of Posttest-then-Retrospective-Pretest Designs

James P. Marshall

Cooperative Extension Service
University of Arkansas Division of Agriculture
jpmarshall@uaex.edu

Brian J. Higginbotham

Utah State University

Victor W. Harris

Utah State University

Thomas R. Lee

Utah State University



Assessing Program Outcomes: Rationale and Benefits of Posttest-then-Retrospective-Pretest Designs

James P. Marshall
University of Arkansas Division of Agriculture

Brian J. Higginbotham, Victor W. Harris and Thomas R. Lee
Utah State University

Abstract: The importance of program evaluation for decision making, accountability, and sustainability is examined in this article. Pros and cons of traditional pretest-posttest and posttest-then-retrospective-pretest methodologies are discussed. A case study of Utah's 4-H mentoring program using a posttest-then-retrospective-pretest design is presented. Furthermore, it is argued that the posttest-then-retrospective-pretest design is a valid, efficient, and cost-effective way to assess program outcomes and impacts.

Need for Program Evaluation

Evaluation is "the systematic collection and analysis of program-related data that can be used to understand how a program delivers services and/or what the consequences of the services are for participants" (Weiss & Jacobs, 1988, p. 49). Although there are a variety of reasons to conduct program evaluation, three of the primary reasons outlined by Little, Dupree, and Deich (2002) involve the ability to:

- (a) make management decisions,
- (b) demonstrate accountability,
- (c) build a case for sustainability.

Prevention programs are becoming more common in many cities across the United States, and at the same time funding agencies are requiring program evaluations to document the effectiveness of funded programs. However, many agencies and organizations that provide prevention programs, such as youth mentoring, are not equipped for or familiar with formal evaluation processes. Many agencies are excellent at providing services but they may not be as

skilled at evaluating the outcomes of the services they provide. The posttest-then-retrospective-pretest research design will enable many smaller organizations, or those organizations with limited experience with outcome evaluation, to efficiently and cost-effectively measure their program outcomes and impacts.

One of the problems associated with many youth mentoring programs is their inability to document required outcomes. If financial support for prevention programs for at-risk youth is to continue, programs must develop evaluation strategies that objectively document program outcomes and impacts (Baldwin, 2000).

Research and Evaluation Strategies

Measuring Change Using the Traditional Pretest-Posttest Method

One of the most respected methods to measure change in individuals is the experimental pretest-posttest design using a control or comparison group (Campbell & Stanley, 1966; Kaplan, 2004). Two reasons for the deference to the pretest-posttest method are its presumed tight scientific control over threats to internal validity and the fact that it can be used to make comparisons between the same people, or groups of people, at different points in time.

However, like all research designs, the pretest-posttest design has some limitations. Limitations may include the difficulty, or impossibility, of locating and maintaining an adequate comparison group. As is the case with many community-based programs, some organizations simply lack the time and resources necessary to conduct pretest-posttest evaluations (Brooks & Gersh, 1998). Further, for pretest-posttest comparisons to be meaningful, participants must be present when the program begins and ends, yet attrition and sporadic attendance are common among community education programs (Pratt, McGuigan, & Katzev, 2000). Another important limitation is that even when complete pretest-posttest information is obtained, actual changes in attitudes, behaviors, or skills may not be evidenced if participants overestimate their attitudes, behaviors, or skills on the pretest (Howard, 1980).

Overestimation on a pretest is likely if participants do not have a clear understanding of the attitudes, behaviors, or skills a program is targeting (Pratt, McGuigan, & Katzev, 2000). Often, it is the participant's lack of knowledge or performance in certain areas that warrant a program intervention in the first place. Participating in the program may show participants they actually knew much less than they originally thought when they completed the pretest. When this is the case, pretest-posttest comparisons are misleading because participants have a different frame of reference after participating in the program than they did before (Howard, et al, 1979). Howard and Daily (1979) were the first to refer to this change in an individual's frame of reference due to program participation as "response shift bias." Simply put, response shift bias can be defined as "a program-produced change in the participants' understanding of the construct being measured" (Pratt, McGuigan, & Katzev, 2000, p. 342).

The following is an example of the misleading effects of response shift bias. A program is developed to teach youth to improve their listening skills. On the pretest they are asked if they actively listen to others when others are speaking. The measurement scale ranges from 1 (never) to 5 (always). One youth perceives herself as someone who usually listens to others and she scores herself at a 4 ("not always but usually"). For the next four months she learns about listening skills and how to actively listen. At the end of the program she realizes that although she has begun using many of the skills she has learned and is a much better listener than before, she is still not a master listener. She now takes the posttest and scores herself at a

4 ("not always but usually"). Her pretest score was 4 and her post test score is 4. In a pretest-posttest design it would appear that her listening skills did not change and that the program was ineffective, when in reality the program probably was effective. What changed was her point of reference. If this youth could re-take the pretest, perhaps she would rate herself differently; however, in a traditional pretest-posttest design this is not an option.

Measuring Change Using the Posttest-then-Retrospective-Pretest Method

The posttest-then-retrospective-pretest research design was created in the late 1970s as a way to control response shift bias in the traditional pre-post design (Howard, Schmeck, & Bray 1979). The post-then-pre design is a way to assess learners' self-reported changes in knowledge, skills, confidence, attitudes, or behaviors (Klatt & Taylor-Powell, 2005a) and it avoids the pretest sensitivity and response shift bias associated with pretest overestimation or underestimation (Howard, 1980; Pratt, McGuigan, & Katzev, 2000; Rockwell & Kohn, 1989).

In the posttest-then-retrospective-pretest design, both before and after information is collected at the same time. The procedures for administering the posttest-then-retrospective-pretest are as follows. At the conclusion of the intervention or program, participants are asked to rate their current levels of knowledge, skills, attitudes, or behaviors. They are then asked to reflect back and rate their levels of knowledge, skills, attitudes, or behaviors prior to participating in the program. By taking the posttest and the pretest at the same time it is more likely that both ratings will be made from the same frame of reference, thus eliminating the effects of response shift bias.

A Case-study: Utah's 4-H Mentoring Program

Utah's 4-H Youth and Families with Promise mentoring program has had great success in using the posttest-then-retrospective-pretest method to measure the program's impact on youth and their parents. At the end of each program (academic) year, every youth who has been enrolled in the program for at least six months and one of his or her parents/guardians are invited to complete a posttest-then-retrospective-pre survey. The self-report survey asks both the youth and his or her parent/guardian to assess their perceptions of how well the youth is functioning in the areas of academic achievement, social competency, family bonds, and delinquent behaviors. Paired-samples *t*-tests are used to compare retrospective pretest scores with the corresponding posttest scores for both youth and parents. Table 1 shows results from the 2004-2005 program year.

Table 1

Paired-samples *t*-test results of youth and parents perceptions of academic achievement, social competency, family bonds, and delinquent behavior.

Variables of interest	Posttest mean score (<i>SD</i>)	Pretest mean score (<i>SD</i>)	Mean change (<i>SD</i>)	<i>t</i>	<i>p</i>
Academic Achievement					
Youth report <i>N</i> =181	24.26 (4.16)	21.67 (5.20)	2.59 (3.72)	9.36	.001*
Parent report <i>N</i> =160	23.17 (5.00)	20.59 (5.38)	2.58 (3.30)	9.86	.001*
Social Competency					
Youth report <i>N</i> =184	32.23 (5.60)	29.30 (6.31)	2.93 (4.42)	9.00	.001*
Parent report <i>N</i> =159	30.31 (5.65)	27.19 (5.84)	3.12 (4.21)	9.35	.001*
Family Bonds					
Youth report <i>N</i> =172	44.88 (7.72)	42.08 (9.15)	2.80 (4.46)	8.20	.001*
Parent report <i>N</i> =157	45.32 (6.88)	42.50 (7.32)	2.83 (5.23)	6.77	.001*
Delinquent Behavior					
Youth report <i>N</i> =178	42.33 (3.87)	43.13 (2.73)	.80 (2.15)	-4.98	.001*
Parent report <i>N</i> =155	42.27 (3.38)	43.09 (2.55)	.82 (1.76)	-5.78	.001*

Both youth and parents reported statistically significant ($p < .001$, two-tailed) improvements in youth levels of academic achievement, social competency, family bonds, and delinquent behaviors.

Benefits of the Posttest-then-Retrospective-Pretest Method

Although the posttest-then-retrospective-pre design is not free from limitations (e.g., accuracy of participant recall and socially desirable responses), it is a valid, efficient, and cost-effective way to assess program outcomes and impacts (Klatt & Taylor-Powell 2005a; 2005b). The post-then-pre design is a simple, convenient, and expeditious method of assessing self-reported behavioral and attitudinal changes in youth and family programming. It is convenient because it is only administered a single time. Only collecting outcome data at the end of a program conserves time and resources and it requires less complicated data management than traditional pretest-posttest designs. The post-then-pre design is also extremely flexible because survey questions can be designed to reflect actual program content, as it may evolve during the course of a program. Finally, research has shown that a post-then-pre design reduces or eliminates response shift bias (Howard, 1980). Although a youth mentoring program example was provided here, the methodology can be adapted and easily applied to other youth and family programs.

References

- Baldwin, C. K. (2000). Theory, program, and outcomes: Assessing the challenges of evaluating at-risk youth recreation programs. *Journal of Park and Recreation Administration, 18*, 19-33.
- Brooks, L., & Gersh, T. L. (1998). Assessing the impact of diversity initiatives using the retrospective pretest design. *Journal of College Student Development, 34*, 383-385.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Howard, G. S. (1980). Response-shift bias a problem in evaluating interventions with pre/post self-reports. *Evaluation Review, 4*, 93-106.
- Howard, G. S., & Daily, P. R. (1979). Response-shift Bias: a source of contamination of self-report measures. *Journal of Applied Psychology, 64*, 144-150.
- Howard, G. S., Ralph, K. M., Gulanick, N. A., Maxwell, S. E., Nance, D., & Gerber, S. L. (1979). Internal validity in pretest-posttest self-report evaluations and the re-evaluation of retrospective pretests. *Applied Psychological Measurement, 3*, 1-23.
- Howard, G. S., Schmeck, R. R., & Bray J. H. (1979). Internal invalidity in studies employing self-report instruments: A suggested remedy. *Journal of Educational Measurement, 16*, 129-135.
- Kaplan, D. (Ed.). (2004). *The Sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, CA: Sage.
- Klatt, J., & Taylor-Powell, E. (2005a). Quick Tips #27: Using the retrospective post-then pre design. Madison, WI: University of Wisconsin – Extension.
- Klatt, J., & Taylor-Powell, E. (2005b). *Quick Tips #28: Designing a Retrospective Post-then-Pre Question*. Madison, WI: University of Wisconsin – Extension.
- Little, P., Dupree, S., & Deich, S. (2002). *Documenting progress and demonstrating results: Evaluating local out-of-school time programs*. Cambridge, MA: Harvard Family Research Project and Washington D.C.: The Finance Project.
- Pratt, C. C., McGuigan, W. M., & Katzev, A. R. (2000). Measuring program outcomes: Using retrospective pretest methodology. *American Journal of Evaluation, 21*, 341-349.
- Rockwell, S. K., & Kohn, H. (1989). Post-then-pre evaluation: Measuring behavior change more accurately. *Journal of Extension, 27*, <http://www.joe.org/joe/1989summer/a5.html>.
- Weiss, H. B., & Jacobs, F. H. (1988). *Evaluating family programs*. New York, NY: Hawthorne.