



**The Challenges of
Implementing Experimental
and Quasi-Experimental
Designs in Teaching
American History (TAH) Grant
Evaluations**

Kimberly S. Cowley
Kristine L. Chadwick
Georgia K. Hughes-Webb
Akisha R. Jones
Susan Voelkel
Damayanti Samanta
Chandra O'Connor

Paper presented at the 2008 Annual
Meeting of the American Evaluation
Association

Denver, Colorado

Edvantia is a nonprofit corporation, founded in 1966, that partners with education agencies, foundations, publishers, and service providers to transform education through research and development. Edvantia provides clients with a range of services, including research, evaluation, professional development, and technical assistance.

For information about Edvantia research, products, or services, contact



P.O. Box 1348, Charleston, WV 25325 • 304.347.0400 • 800.624.9120 • fax 304.347.0487
One Vantage Way, Suite D-210, Nashville, TN 37228 • 615.565.0101 • fax 615.565.0112
info@edvantia.org • www.edvantia.org

© 2008 by Edvantia, Inc.

Reproduction of this document is permitted with the following citation:

Cowley, K. S., Chadwick, K. L., Hughes-Webb, G. K., Jones, A. R., Voelkel, S., Samanta, D., & O'Connor, C. (2008). *The challenges of implementing experimental and quasi-experimental designs in Teaching American History (TAH) grant evaluations*. Charleston, WV: Edvantia.

This publication was prepared by Edvantia staff for presentation at the 2008 Annual Meeting of the American Evaluation Association. Its contents do not necessarily reflect the positions or policies of the Teaching American History grant recipients cited within the document or Edvantia.

Edvantia is an equal employment opportunity/affirmative action employer.

Contents

Introduction	1
Challenges and Practices	2
Design Changes.....	3
Recruiting Comparison Groups	5
Collecting Data from Comparison Groups	9
Summary.....	11
References.....	12

Introduction

Teaching American History (TAH) is a discretionary grant program funded under the Elementary and Secondary Education Act (ESEA). The program supports projects focused on raising student achievement by improving teachers' knowledge, understanding, and appreciation of U.S. history (U.S. Department of Education, 2007). Local education agencies compete for the grants and must agree to carry out the proposed activities in partnership with one or more institutions or organizations that have extensive content expertise, such as institutions of higher education, nonprofit history or humanities organizations, libraries, and museums. The grants are used to provide extensive professional development for American history teachers to improve the quality of instruction. "The goal of this program is to demonstrate how school districts and institutions with expertise in American history can collaborate over a 3-year period to ensure that teachers develop the knowledge and skills necessary to teach traditional American history in an exciting and engaging way" (U.S. Department of Education, 2008).

As a requisite component of funded TAH projects, grantees must now include two GPRA (Government Performance and Reports Act) performance measures—one focused on teachers and one on students. The student measure requires either experimental or quasi-experimental evaluation designs. These two performance measures are as follows:

1. Students in experimental and quasi-experimental studies of educational effectiveness of Teaching American History projects will demonstrate higher achievement on course content measures and/or statewide U.S. history assessments than students in control and comparison groups.
2. Teachers will demonstrate an increased understanding of American history through the use of nationally validated tests of American history that can be directly linked to their participation in the Teaching American History program.

To date, Edvantia has served or is serving as the external evaluator for eight TAH grant recipients, beginning with the first year of the TAH grants in 2003:

1. *Connections to the Past*, a 3-year Virginia project involving Spotsylvania County, Stafford County, and Fredericksburg City Schools (2003-2006)
2. *People and Moments in Time*, a 3-year Virginia project involving Spotsylvania, Stafford, and Caroline County Schools (2005-2008)
3. *Evaluation of Project TEACH II: The Constitution in Historical Context*, a 3-year West Virginia project involving Regional Education Service Agency III (2006-2009)

4. *American HEART: Americans Who Made History*, a 3-year statewide West Virginia project involving teachers from 13 counties (2005-2008)
5. *Forging Democracy: Change, Conflict, and Continuity*, a 3-year Virginia project involving Spotylsylvania, Stafford, and Caroline County Schools (2007-2010)
6. *Teaching American History in Miami-Dade*, a 3-year Florida project involving Miami-Dade County Schools (2007-2010)
7. *American Origins: Hidden Histories in Our Midst*, a 3-year Virginia project involving Charlotte, Danville, Halifax, and Pittsylvania County Schools (2007-2010)
8. *Shaping and Reshaping America 1787-1920*, a 3-year Pennsylvania project involving Philadelphia City Schools (2007-2011); Edvantia to serve as the evaluator for the third year and for the 1-year extension

Seven of these evaluations were designed as quasi-experiments, and one evaluation was designed as a true experiment (Shadish, Cook, & Campbell, 2002).

Across the eight TAH evaluation projects, Edvantia evaluators have experienced a variety of challenges as they attempted to implement experimental or quasi-experimental evaluation designs on relatively small budgets. Working with slightly different designs and methods, these evaluators have learned a few lessons about how to adhere to the rigorous evaluation mandate in the project solicitation, soothe client fears, partner with clients to implement evaluation designs, collect data for both formative and summative purposes, stay within budget, and manage at least a few threats to validity. Three of the main challenges include (1) revising evaluation designs at the initial design phase or during implementation from true experiments to quasi-experiments with delayed treatment; (2) negotiating and recruiting teacher comparison groups; and (3) collecting data from comparison group teachers. This paper details these challenges and offers examples of how the evaluators attempted to overcome them.

Challenges and Practices

Throughout the eight TAH evaluations, a number of challenges have emerged throughout planning and implementation phases. We have categorized the majority of these challenges, and our resulting practices, into three categories: design changes, recruiting comparison groups, and collecting data from comparison groups.

Design Changes

Across the evaluation projects, several types of design changes took place before and/or during the project evaluation. These changes are described below.

Design changes before implementation. As proposals are being developed and finalized, evaluators and project staff may discuss and negotiate alternate research and evaluation designs before settling on one that best suits their needs.

- *Example:* Edvantia evaluators were working with current TAH clients to submit a new TAH proposal. The project currently under way incorporated a true experimental design; clients had agreed to this design because competitive priority points were being appointed by the U.S. Department of Education reviewers for proposals with this component. However, during the new funding cycle, no extra points were being awarded for experimental designs, and the clients opted instead for a quasi-experimental design.
- *Example:* Another possible scenario is that faced when proposal writers are not the subsequent project directors of funded grants. As a result, project priorities, activities, and/or goals may change, necessitating revisions to the evaluation plan.

Design changes during implementation. Even after proposals are funded and implemented, changing contexts may sometimes necessitate modifications to the planned evaluation design. For example, a project may be changed from an experimental design to a quasi-experimental delayed treatment during the final year of a project. By its very nature, the experimental design requires that some applicants be randomly assigned to the control condition, thereby making them ineligible to participate in the project treatment. Because those who volunteer to participate in a project generally do so because they want to receive the treatment, those individuals assigned to the control condition are likely to be disappointed (e.g., experience resentful demoralization) or to seek similar services elsewhere (e.g., develop compensatory rivalry); both situations present possible threats to construct validity (Brewer, 2000; Shadish, Cook, & Campbell, 2002). Switching to a quasi-experimental delayed treatment design during the final year of the project allows controls from earlier years to be assigned to the treatment condition in the final year. This can be an effective strategy/incentive to help ensure adequate applications during earlier years and has the added benefit of providing built-in comparison data for the final-year participants.

- *Example:* The proposal for one TAH project called for an experimental design. Random assignment of treatment and control teachers from the applicant pool was carried out in Years 1 and 2. However, although the applicants had been informed that they might be assigned to the control group, many of those assigned as controls in Year 1 were very disappointed; especially frustrated were those who reapplied in Year 2 and were again randomly assigned to the control group. These individuals were not mollified by a stipend for their services as control teachers; they wanted to

be project “Fellows” who went on field trips to Philadelphia, Gettysburg, Williamsburg, and various other historical sites; received numerous children’s history books and other resources to use in their classrooms; and received three graduate credits if they completed all of their work. This disappointment and frustration led to much grumbling and frustration, reduced compliance with data collection, and lessened the likelihood that sufficient numbers of teachers would apply to the project in the third and final year. Consequently, in Year 2, applicants were told that if they were randomly assigned to the control group in Year 2, they would become project participants in Year 3, if they so desired. Of the 43 randomly assigned control teachers in Year 2, only 3 (7%) did not fulfill all control obligations during the year.

The decision to switch to a quasi-experimental delayed treatment for Year 3 necessitated several adjustments. First, the Year 3 treatment group did not require another comparison group, since their pretest scores from the prior year could be used as the control condition. Second, pre/post changes in their students’ knowledge test scores were compared across the 2 years. And, third, a “control construct” or “nonequivalent dependent variable” was added to strengthen the evaluation design. As elaborated by McKillip (1992), the control construct design involves testing or observing the group of participants (teachers) for multiple constructs: one construct targeted by the intervention (in this case, American history) and one or more similar constructs not subject to the intervention (in this case, geography). The geography knowledge test covered material similar to U.S. history; presumably would be unaffected by the TAH project treatment; and, therefore, in the absence of a concurrent comparison group, would control for the influence of such threats to internal validity as history- or test-wiseness. For all Year 3 participants, the pre/post change in their history knowledge was compared to the pre/post change in their geography knowledge.

In other circumstances, even quasi-experimental designs may need to be modified during project implementation. In the example described below, a quasi-experimental design involving treatment and comparison group teachers was modified to a one-group design (no comparison group) due to circumstances leading to the loss of the comparison group.

- *Example:* One TAH project had a small treatment group—20 teachers who participated for all 3 years of the project and who were located in multiple counties throughout a state. Project staff attempted to recruit comparison teachers, matched to participants by school or district, but did not complete this recruitment until the second year of the project. In Year 2, 14 comparison teachers completed the same instruments as the participating teachers but did not provide the necessary student data to complete the student GPRA measure, despite numerous mailings and e-mailings. Because the teacher GPRA measure did not explicitly require a comparison group, and none of the project measures specified a comparison group, and the comparison teachers were difficult to contact and noncompliant in providing

necessary student data, the comparison group was dropped from the Year 3 evaluation plan. However, the participants' 3-year involvement in the project provided longitudinal data to demonstrate their growth across the length of the project, and their students' course grades were compared for Years 2 and 3.

Recruiting Comparison Groups

Effective TAH evaluations require the cooperation of participating teachers and, a majority of the time, a group of comparison teachers. Recruiting teacher comparison groups proved challenging in most of the eight TAH projects. Negotiating recruitment strategies and securing participation from comparison teachers required different solutions in every project.

Involvement of the project director and/or staff. The active role of the project director has been critical in evaluators' efforts to recruit comparison teachers to participate in evaluation activities. Project directors often have day-to-day contact with district or school personnel and teachers throughout the area served by the project and often have influence that is difficult for evaluators to obtain. They are able to create buy-in to evaluation procedures and exert strong pressure to comply amongst all participants when strong relationships have been developed. Project managers have also been influential in creating buy-in among participating teachers, which has increased their willingness to participate in project activities. Increasing buy-in from project staff, participating teachers, and comparison teachers can become a form of action research wherein the evaluator and project staff work together to implement as rigorous an evaluation design as possible (Owen, 2007).

- *Example:* In most of the TAH evaluation projects, geographical distance and accompanying travel expenses prohibit much on-site involvement by evaluators. Having project directors or key staff "actively" supporting evaluation activities has been critical. Project directors/key staff are often staff members from participating schools or districts, people who already have working relationships at the schools. As a result, they have been most effective in talking to school staff, explaining the need for and responsibilities of control/comparison teachers, and following up on-site throughout the project year as needed to support data collection efforts. Note: a precursor to having project directors/key staff fully engaged in recruiting comparison groups is making sure they fully understand the evaluation design and methods, as well as the importance (and potential benefits) of gathering quality data.
- *Example:* Another way in which evaluators have secured staff involvement in recruiting comparison teachers is by having the participating (treatment) teachers themselves recruit their comparison counterparts. During the first year of one project, evaluators had designed a sophisticated method for matching comparison teachers to participating teachers using locale, grade level taught, and years of experience as key factors; a pool of comparison candidates was formed by using all of the history or social studies teachers in participating teachers' schools. However,

when the selected comparison teachers were contacted regarding their selection and requested participation, most refused either actively or passively. Response rates to requests for data were abysmal. Evaluators and project leaders worked collaboratively to determine what recruitment strategies might work best within the context of the multicounty evaluation. During the second year, participating teachers were asked to select from among their colleagues a teacher to serve as their comparison. The greatly improved response to this method of comparison teacher recruitment resulted in successful data collection and meaningful comparisons. Although the years of experience may have differed slightly between participants and their comparisons, the locale and grade level factors were fairly good matches because the colleagues taught in the same school buildings.

The above examples—with their emphasis on the importance of strong relationships between and among evaluators, project staff, participating teachers, and comparison teachers—align with the approach of participatory evaluation. This model advocates joint ownership and control of evaluation decision making by those stakeholders most closely connected with and affected by the program (Cousins & Earl, 1995; Cousins & Whitmore, 1998).

Varying criteria for selection of comparison teachers. Project evaluators have encountered a variety of challenges in their attempts to recruit teachers for evaluation activities. These challenges have been dependent upon the specific data collection procedures that are required in each evaluation design.

- *Example:* Most of the TAH evaluations call for a pre/post knowledge test for treatment and control/comparison teachers. Some projects were able to identify and recruit comparison teachers with characteristics similar to those of the treatment teachers (i.e., demographic variables such as education level and years of teaching experience). In other projects, in which it was more difficult to find comparison teachers, statistical techniques (ANCOVA) were used to control for these variables in subsequent analyses.
- *Example:* As noted earlier, current TAH projects are required to include a quasi-experimental or experimental design so that evaluators can compare the achievement of treatment teachers' students to their comparison counterparts. The need to gather student data has implications for the selection of comparison teachers. To achieve reasonably equivalent groups of students, project evaluators have attempted to focus on school-level as well as teacher-level demographics. Such school-level variables include indicators such as grade level, Title I status, AYP status, and/or geographical locale.

Incentives. One of the most notable challenges in recruiting teachers for evaluation activities is the ability to provide fair and just incentives to both participating and comparison teachers, although the challenge is particularly pertinent in recruiting comparison teachers. Evaluations that have had the least participation among comparison teachers are those that

failed to offer those teachers any substantial incentive to participate. Conducting evaluations with limited funds often restricts the capacity of project staff and evaluators to provide teachers with attractive incentives. Without sufficient financial (or other) incentive for participation, teachers have been very reluctant to commit their time to evaluation procedures.

- *Example:* Quite a few of the TAH projects now include some type of incentive to be provided through project (not evaluation) funds to ensure participation of comparison teachers. Such participation differs depending on the particular evaluation design. For instance, one of the earliest projects did not include direct student testing, so comparison teachers had only to complete a pre/post knowledge test during the school year. In order to secure their cooperation, project directors administered the pre/post tests at a local restaurant, providing free refreshments while comparison teachers took the assessment. This not only provided an inviting, relaxing environment but also helped standardize testing conditions.
- *Example:* In projects in which student pre/post testing was a required element, comparison teachers were committing to a much more time- and labor-intensive set of activities. Not only were they to complete their own pre/post testing, but they were also required to administer such assessments in their classrooms with their students after securing parental permission for such testing. For several of these projects, project directors actually paid comparison teachers a stipend for their participation in these data collection activities. Such monetary incentives ranged in value depending on the size of the project. These incentives must be built into the proposal budget at the time the proposal is submitted; otherwise, it would be difficult to carve out sufficient funding after project activities are fully planned.

A resulting challenge from the use of monetary incentives is the need to make sure that participating (treatment) teachers understand that their incentive is the project treatment itself (free professional development), and that they would not receive such monetary incentives. Although this might seem fairly self-evident, several projects did encounter this resentment, especially those projects in which treatment teachers selected their own comparison counterparts. This is another instance in which it is helpful to have project directors firmly in support of the evaluation; they can then reiterate the message directly to treatment teachers that the monetary incentive is merely a way to compensate the comparison teachers for agreeing to participate in data collection activities which do not benefit them in any other way.

“Contamination.” Although the use of comparison groups within participating schools and/or districts has proven successful, there is the possible challenge of contamination—comparison teachers’ coming into contact with materials or strategies from the project treatment. This is especially true in those situations in which participating teachers choose their own comparison. Although such a threat exists any time project participants and comparisons teach within the same school, this threat is offset by the increased likelihood of involving highly comparable pools of students in the study. We’ve given explicit instructions to participants to

avoid this, and then asked both sides, in the final data collection, whether such “sharing” had occurred.

- *Example:* This has been an issue faced in all of the TAH projects involving comparison or control groups. Evaluators have worked closely with project directors to impress upon them the importance of confining the treatment to participating teachers and their classrooms during the evaluation period. We have worked to make sure this message is received and understood at the participant level. Although this practice seems exclusionary and out of character with the goal of improving education for all students, it is imperative to be able to determine project impact that comparison/control teachers do not receive project knowledge, strategies, artifacts, and so forth. To determine to what degree such contamination might have occurred within projects, questions are added to the post knowledge test for both groups to determine whether (a) treatment teachers have shared knowledge/strategies/information with anyone besides other participating teachers and (b) whether comparison teachers have received any project materials.

Random assignment. Evaluators also encountered challenges with the random assignment activity within the one experimental project.

- *Example:* For the first 2 years of this project, half of the applicants each year were assigned to the treatment group, and the other half to the control group. During the first year, when this randomization process was new to the project staff, and when the application process itself was a bit chaotic due to the newness of the grant and the delay in award notification, there was a bit of a learning curve on how this randomization had to take place. The evaluators had to field a number of requests for “exceptions” to the randomization process; eventually, project staff came to understand that people could not be moved arbitrarily from one group to another.

In the second year, a new challenge emerged when two married couples applied to the project and the random assignment split both couples into the two conditions (treatment and control). For evaluation purposes, this situation was not desirable because of the high likelihood of contamination (e.g., the spouse in the treatment group sharing information and materials with his or her spouse in the control group). To keep both couples from dropping out of the project, and in consultation with the federal technical assistance providers, the evaluators and project staff agreed that the two “control” spouses could participate in all project activities. However, their data were excluded from the experimental analyses related to the evaluation, and they were classified as control “dropouts.”

Finding sufficient number of comparisons. Another challenge faced in some of the evaluations was finding a sufficient number of teachers willing to serve as comparison teachers. This scarcity could be due to a number of reasons; for instance, limited awareness and/or confusion about comparison group obligations, oversaturation of TAH exposure from previous

projects within a school or district, and/or scheduling of comparison activities (i.e., teacher and student pre/post testing occurred traditionally at the beginning and end of the school year).

- *Example:* In one multi-district area, in which multiple TAH projects had been funded, it became exceedingly more difficult to find teachers within each participating school who had not been exposed to previous TAH treatment. As a result, the pool was broadened to include all schools within the participating districts, rather than just within those participating schools. The pool was further expanded to allow for interdistrict identification of comparison teachers, when necessary.

Collecting Data from Comparison Groups

Collecting high-quality data from treatment and comparison groups is crucial for experimental and quasi-experimental TAH evaluations. It is often relatively easy to collect data from treatment teachers who have agreed to participate in all aspects of the TAH project, including the evaluation. However, collecting data from comparison groups has required evaluator flexibility and extensive use of persuasion skills, thus highlighting both the technical and people skills that evaluators must have to be effective.

Educating project staff about the importance of good data collection. Some of the same challenges that relate to teacher recruitment for evaluation activities overlap with the challenges evaluators face with collecting data from comparison group participants. One main issue with comparisons is that they may consent to be in the comparison group, yet not comply with data collection procedures.

- *Example:* Successful data collection relies heavily on the project directors and/or key staff, who can request, nudge, or even more forcefully ask that comparison teachers comply with data collection activities. This often requires patience on behalf of evaluators as they help project staff understand evaluation in general, so that project staff can communicate the importance of data collection to both comparison and treatment groups. For all projects, evaluators have strived to translate our language and standards into a format that our clients and other stakeholders (including comparison teachers) will understand, buy into, and actively support. Helping all stakeholders understand that accurate and complete data collection is key to ensuring study validity and to helping reveal project impact is a necessity for each and every project.

On a related note, personal contact is critical. If geographic distance and travel expenses prohibit evaluator on-site presence (as is the case with most of our TAH evaluations), then project directors and/or key staff (and even participating teachers) can provide personal contact that helps foster participation. Evaluators must be sure to provide project staff with the proper message, and then follow up with electronic communications if needed. (Note: this personal contact is critical for sustaining active participation by treatment teachers as well.)

Project director's influence. The influence or “clout” of the project director has also been critical to successful data collection. Usually, evaluators have no direct influence over the comparison teachers.

- *Example:* Some project directors have more clout than others. In one project, a director is also the social studies director of the district, has been there for a number of years, and subsequently has established good relationships with the history and social studies teachers. Therefore, she was able to use her personal relationships to persuade comparison teachers to comply with all requested activities.
- *Example:* In each project, participating teachers receive substantial benefits—free field trips, classroom materials, graduate credit, etc. Although participation in evaluation-related data collection activities cannot be mandated by evaluators, the project director has more of an opportunity to explicate/expect full compliance when teachers agree to participate as treatment or control teachers. Some projects include language in the agreements to indicate that certain activities/benefits could be withheld if participants do not comply with all data collection activities.

Trust. Evaluators must also deal with the challenge of trust—some teachers (or other stakeholders) may be distrustful of evaluation, thinking data might be used against them in some way. Again, the reputation of the project director can help ease any anxieties.

- *Example:* In several of our projects, the same staff have served as the directors for several TAH projects. As a result, teachers in the districts know and trust them, and they have seen firsthand that there has been no negative “fallout” from other teachers who participate in data collection activities. Too, for multiyear projects or districts with several projects, the projects themselves may have acquired a “good reputation,” resulting in teachers being eager to participate in the treatment or at least feeling safe enough to participate as a comparison. For those projects in which participating teachers recruit their own comparisons, trust has been much easier to establish because the teachers involved already know and work with one another.

Incentives. As noted earlier, incentives may increase the likelihood that teachers will agree to serve as comparisons. Incentives may also increase the likelihood that comparison teachers will fully complete all associated data collection activities.

- *Example:* The TAH project that seems to have been most successful with financial stipends linked the stipend vouchers to the final data collection activity. In other words, comparison teachers must have completed the final activity before they received their stipend. Another project paid part of the stipend at the beginning of the project year and the remainder at the conclusion of the project year. Still another project held the final data collection activity for comparisons at a restaurant. Not only did comparison teachers receive free refreshments; the arrangement also provided a controlled setting for the posttest administration.

Finally, one project provided controls from the first 2 years of the project with the incentive/assurance that they would become treatment teachers in the third and final year.

Comparison teachers without students or with students from different grades. The last challenges we encountered are associated with collecting student data from comparison teachers. We have faced situations in which comparison teachers did not have classroom students and where comparison teachers do not teach the targeted grade level.

- *Example:* In some projects, several of the comparison teachers (and occasionally a participating teacher) ended up not having any students to complete the pre/post knowledge test. In most cases, these teachers held resource or special education positions, and did not have their own classrooms. Given the difficulties encountered in finding an adequate number of comparisons for several of the projects, these teachers were retained as comparisons and their own pre/post knowledge test results were used, although no student data were available for these teachers.
- *Example:* In one project in which the treatment year extends across school years, and the treatment is focused specifically on one grade level, we have encountered situations in which teachers (both treatment and comparison) were reassigned to different grade levels from one school year to the next. As a result, their students are no longer in the appropriate grade level. In these instances, we do not collect student data.

Summary

It is not in any way original to say that carrying out experimental and quasi-experimental field studies is not easy. Between the long list of threats to validity, the short list of clients and participants who want to be involved in experiments, and the GPRA measure that demands comparison groups of students, the evaluator is sandwiched between that proverbial rock and hard place. Satisfying the formative *and* summative evaluation needs of a client further complicates the endeavor.

So what is an evaluator to do when faced with design demands, budget constraints, and, often, client ambivalence or lack of appreciation for the mandated design? Einstein is credited as saying that an experiment is something everyone believes except the person who did it. Given what we evaluators know about what happens when an elegant design is taken into the field, we can appreciate Einstein's words of wisdom. We do the best we can. We learn from our colleagues' successes and failures.

In the eight projects we have evaluated, our staff have been fairly successful at tailoring the evaluations to the specific challenges of each project while also aligning the evaluation with federal reporting requirements. Custom designs are the norm in evaluation, and definitely so

for evaluators employing a utilization-focused perspective. But being good researchers is not enough; our effectiveness also depends on effective communication. Having the accompanying “people” skills to be able to translate designs into nontechnical, nonthreatening language for project staff and participants helped to foster understanding and “buy-in” for associated data collection activities. Also, working closely with project staff in identifying, recruiting, and maintaining the cooperation of comparison-group teachers helped to ensure adequate comparison data in most projects. We hope that the challenges discussed in this paper, as well as the examples provided, will help other evaluators gain insights into their own work involving experimental and quasi-experimental research designs.

References

- Brewer, M. (2000). Research design and issues of validity. In H. Reis and C. Judd (Eds.), *Handbook of research methods in social and personality psychology*. Cambridge: Cambridge University Press.
- Cousins, J. B., & Earl, L. M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation, New Directions for Evaluation, 20*, 5-23.
- McKillip, J. (1992). Research without control groups: A control construct design. In F. B. Bryant, J. Edwards, R. S. Tindale, E. J. Posavac, L. Heath, E. Henderson-King, & Y. Suarez-Balcazar (Eds.), *Social psychological applications to social issues: Vol. 2. Methodological issues in applied social psychology* (pp. 159-175). New York: Plenum Press.
- Owen, J. M. (2007). *Program evaluation: Forms and approaches (3rd ed.)*. New York: Gullford Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.
- U.S. Department of Education. (2008). *Teaching American History grant program: Purpose*. Retrieved October 16, 2008, from <http://www.ed.gov/programs/teachinghistory/index.html>