



### **SAMPLE MATERIAL: Recruiting Comparison Groups**

**Purpose:** Participation of comparison school staff, teachers, and students is essential to successful implementation of a rigorous, quasi-experimental evaluation. However, these groups are often reluctant to be part of research studies because they don't believe they'll benefit from participating. This reading offers strategies to consider for effectively involving these important players.

**Source:** Excerpted from a research paper presented by Edvantia at the 2008 Annual Meeting of the American Evaluation Association: 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.



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

### 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 to 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.