Peace Works: A Program Evaluation Model

A Paper Presented at the
Annual Meeting of the American Educational Research Association
April 3, 2002
New Orleans, LA

Patrice LeBlanc, Ed.D.
Nova Southeastern University, Ft. Lauderdale, FL

Candace Lacey, Ph.D.
Barry University, Miami Shores, FL
School violence is a pressing problem (Sautter, 1995; U.S. Department of Justice, 2000), and the resulting pressure on schools to implement violence prevention programs is enormous. However, of the more than 80 violence prevention programs on the market, few of them have undergone evaluation studies (Knowles, 2001). This paper describes in detail the model employed to evaluate one such prevention program, “Peace Works,” implemented in the two-year Allegany Grant between the Peace Education Foundation (PEF) and the Miami-Dade County Public Schools (M-DCPS).

The grant project goals, addressed through implementation of “Peace Works,” were to: promote students’ prosocial behavior; create a more caring, supportive school climate; teach parents constructive problem solving and anger management; and encourage parents’ positive affiliation and involvement with the schools. Although the research model was designed to assess this specific project’s goals and outcomes, it has potentially broader use for assessment of other violence prevention programs.

Method

Setting and Participants

The grant project was conducted in Region VI of the M-DCPS, a semi-rural high poverty area with many migrant workers. Eight elementary schools and two middle schools were chosen to participate based on similar demographics and economics, thereby controlling for those characteristics. The elementary school populations ranged from 800-1100 students, while the middle schools had approximately 1300-1500 students. Thus, participating students totaled approximately 10,000.
Once identified, the schools were randomly selected as treatment or control schools. From the treatment schools, two elementary schools and one middle school were randomly selected for further micro-level evaluation along with the treatment middle school.

The “Peace Works” Program

The PEF’s “Peace Works” program teaches students the dispositions, behaviors, and skills necessary to peaceably resolve conflict. The program content consists of six essential components: communication building, rules for fighting fair, understanding conflict, the role of perceptions, anger management, and effective communication. The components are taught through grade-level lessons, student workbooks, teacher’s manuals, and other materials (e.g., “I Care Cat” puppet). For correct implementation, the teacher teaches lessons a minimum of once per week.

Evaluation Questions

Four primary evaluation questions were phrased directly from the project goals to assess the foci of the project, with sub-questions generated to address specific target areas (e.g., discipline and safe school processes). Also, data were collected on the outcomes related to student-centered growth (e.g., student discipline referrals).

Evaluation Design

A mixed method research design was employed using simultaneous methodologies (Tashakkori & Teddlie, 1998). This design was selected because it addressed concerns about validity and reliability, and thus the transferability of the findings to other settings while providing the potential to add to the empirical and descriptive literature on violence prevention program evaluations.
The quantitative methodology, an experimental pre-test and post-test design (Gay & Airasian, 2000), provided the primary data for assessing outcomes related to students’ growth. This design was selected to insure that gains made in students’ prosocial skills and reductions in aggressive behavior were the result of the conflict resolution program and not the result of other factors (e.g., maturation).

Qualitative methodology provided additional data on outcomes, adding depth and breadth that enhances the transferability of findings (Seale, 1999). The qualitative methods of interviews, focus groups, and observation obtained participants’ perceptions, allowing the evaluators to better understand the how and why behind project outcomes (Morgan & Krueger, 1993; Rubin & Rubin, 1995). The qualitative design used triangulation of data sources, collection, and analysis procedures to provide validity (Silverman, 1993, 2000; Tashakkori & Teddlie, 1998). Reliability of the qualitative data was addressed using standard procedures in the field such as defining the reporting of methods, the theoretical framework that guided the study, and using multiple evaluators (Seale, 1999).

Further, data were collected on both macro and micro-levels. Macro-level data allowed a global project perspective, comparing treatment and control schools on a number of school-wide data sources. Micro-level data provided a detail perspective, allowing for treatment and control schools comparisons at the classroom level.

For micro-level comparisons, exemplary teachers were identified at the third grade level in two, randomly selected, micro-level elementary schools and at the two control elementary schools, while one exemplary teacher was identified at the sixth grade level in the treatment and control middle schools. School administrators used a list of exemplary teacher characteristics culled from the literature to identify the participating teachers. As an incentive to maintain
participation over the two-year period, the participating micro-level teachers had the option to earn points toward state re-certification for their work on the project.

Exemplary teachers are more self-actualized and have higher self-efficacy, thus they tend to be more persistent and successful at implementing change programs (Fullan, 2001) and would therefore implement the conflict resolution program. In addition, selection of exemplary teachers helps to control for effects of poor teacher self-efficacy, which has been found to negatively impact the implementation of a conflict resolution program (LeBlanc, Lacey, & Mulder, 1998).

Theoretical Framework

A theoretical framework guided the evaluation, providing a structure for thinking about data and a basis for generalizing out qualitative findings (Seale, 1999). Furthermore, the theoretical framework provided a triadic structure for synthesizing the results into an executive summary format that made for easy reading for the non-educator. This framework was a context, process, and outcomes model (Silverman, 1993).

Context in the research model described here refers to environment, often termed culture or climate. Since changes in the climate of the school are typically intertwined with the implementation of conflict resolution, perceptions of school climate and safety were measured for context.

The second component of the model, process, assessed the use of problem-solving skills and levels of self-responsibility among adult project participants. Identifying the adult participants’ understanding of the content in the conflict resolution program was critical to the accurate delivery of the program to students.
Outcomes, the third component of the research model, were specifically defined in the grant project goals: decreased discipline referrals, reduction in students’ angry and aggressive behavior, students’ development of prosocial skills, and increased parent-school affiliation.

Data Sources and Instruments

Macro-level Data: Treatment and Control Schools. Four sources of data were used at the macro-level. First, data on demographics and economics from the district’s school profiles identified control variables.

Second, referral data from the district’s database were collected to ascertain the influence of the conflict resolution program on students’ discipline referrals. Referral data were used since some literature supports its use as one indicator of program success (Speirs, 1994; Tolan, Keys, Chertok, & Jason, 1990), and school districts use it as an indicator of school-wide discipline effectiveness.

However, some difficulties have been noted in using referral data across schools bringing into question the reliability and validity of such comparisons (LeBlanc, Lacey, & Griffin, 2001). There is variation in school climate, discipline, and administrative styles from school to school. For example, one person may record only serious offenses in a referral database, while another may record every incident. This variation makes comparisons of referral data across schools unreliable. Another example is that referral rates are typically low at the beginning of the year, as students tend to be on their best behavior, then they increase at the end of the school year when students become restless anticipating summer vacation. These variations in student behavior within the school year make comparisons from the beginning of the year to the end invalid. In order to compensate for the aforementioned problems as intervening variables in the evaluation,
school-wide referral data for the entire year were totaled and each school was compared to itself for the two-year grant period.

Additionally, the referral data were reviewed by categories identified in the Allegany Grant: (a) disobedience, (b) disruptive behavior, and (c) disrespectful language. These categories were targeted for reduction since they aligned closely with the skills taught in the conflict resolution curriculum and, hence, were thought to provide a measure of project success. This approach is supported in the literature by the work of Bey and Turner (1996) who propose a continuum of escalating violent behaviors. Thus, classifying behaviors provides insights into the type of antisocial behavior that occurs and can help administrators more fully understand the levels of antisocial behavior and violence at their schools.

Third, for the project’s context or the schools’ climate, each school’s School Climate Surveys over two years were compared. Measuring school climate is important since it is viewed as the context into which a conflict resolution program is entwined. As the program is implemented, there is a collateral development of “a learning climate that is committed to equity and social justice” (Bodine & Crawford, 1998, p. 11). This climate supports and encourages the behaviors of conflict resolution as part of the school culture. Thus, “the community lives by a credo of nonviolence and multicultural appreciation” (Dejong, 1994 as cited in Bodine & Crawford, 1998) and, hence, becomes a peaceable school. Since the School Climate Survey is part of the state of Florida’s requirements for school improvement and accountability and is used annually, it was a readily available tool that could be incorporated into the evaluation.

The surveys were distributed to a 25% random sample of students who carried the surveys home to parents. Students and staff received the surveys at school, again with distribution based on a 25% random sample. Recent reliability scores on the survey using the
Cronbach’s Alpha were .96 for the parent form, .88 for the student form, and .88 for the staff form (Romanik & Froman, 1998).

Finally, parent logs were to be collected from all treatment and control schools to measure parent affiliation (e.g., Title I school contact logs). However, all of the schools that participated in the project had little to no parental involvement as reported by the PEF staff involved in the project and the assistant principals and teachers interviewed by the evaluators. Instead, frequency counts of attendance at training sessions were used as the indicator of parental affiliation, which did not allow for comparisons between treatment and control schools.

**Macro-level Data: Treatment Schools Only.** Conflict resolution training pre-tests and post-tests were to be administered for all training sessions to ascertain teachers’, staff’s, and parents’ understanding of content knowledge. A 20 item pre-test and post-test was developed with high content validity in order to provide appropriate assessment of training outcomes (Fowler, 1998; Posner, 1995). During the first year of the project it was piloted, with intended use in year two of the project, to be administered at the beginning and end of each training session.

Second, through interviews and/or focus groups with teachers and administrators, data were collected on participants’ perceptions of conflict resolution program implementation. Although plans were to initially include parents as a data source, lack of parental involvement in the schools precluded their involvement. In order to enhance the validity of qualitative data, protocols were developed (Rubin & Rubin, 1995) that prompted the participants to express their thoughts. Additional probing questions were asked to further elicit rich descriptions of the participants’ experiences (Silverman, 1993).
**Micro-level Data.** Teachers in grades three and six rated their students using the School Social Behavior Scales (SSBS). The SSBS is a valid and reliable instrument used to measure students’ prosocial skills and antisocial behavior (Merrell, 1993, 2001); therefore, it was used as a pre-test and post-test measure of students’ outcomes comparing treatment and control classes. The teachers conducted the first administration just prior to program implementation, and again at the end of the school year.

A summary of all data sources by method and level is depicted here.

-------------------------------------------------------------------
Insert Figure 1 here.
-------------------------------------------------------------------

**Results**

The results of the evaluations using the model are documented in detail elsewhere (Lacey & LeBlanc, 2001; LeBlanc & Lacey, 2002, LeBlanc & Lacey, 2000). The focus of this paper is how those results connect to the model and its usefulness for assessing the “Peace Works” program. The subsequent paragraphs summarize those connections with Figure 2 providing a graphic organizer.

-------------------------------------------------------------------
Insert Figure 2 here.
-------------------------------------------------------------------

**Model Strengths**

**Mixed Methodology Research Design.** The most useful aspect of the model was the mixed method research design using simultaneous methodologies (Tashakkori & Teddlie, 1998).
It allowed for simultaneous data collection in multiple forms that later became extremely valuable. An example follows.

The results of the SSBS scores for year two indicated that the treatment middle school students had significant decreases in hostile/irritable behavior (p = .050) and significant decreases in antisocial/aggressive behavior (p = .012), with a total antisocial behavior significant decrease (p = .019). The middle school treatment students also had significant increases in self-management (p = .001) and academic skills (p = .000), with significant increases in total social competence (p = .003). On the other hand, the control middle school had one social competency area, interpersonal skills, that reached significance (p = .022). Since the research design used a mixed methodology, the evaluators had collected enough interview data to discover why the control middle school had a significant difference.

Unlike the treatment school, the control school had two school-wide violence prevention related programs: the Comer Model and Peer Mediation. The Comer Model was used in the school for six years prior to the Allegany Grant project. Established by James Comer at Yale University, the Comer Model has a long history of success with high minority, high poverty schools (Comer, n.d.; Ramirez-Smith, 1995; Traub, 1999) like those in the project. Among the many areas targeted for development in Comer Model schools, one ties directly with area of interpersonal skills measured directly by the SSBS: “a model of respectful conversation” that is based on “principles of consensus, collaboration, and no-fault” (Traub, 1999, p. 65). Furthermore, the school ran a peer mediation program that taught conflict resolution skills based on similar concepts to those in the “Peace Works” curriculum. Thus, the students at the control middle school may have felt the long-term impact of the Comer Model on interpersonal skills in the school culture and been involved in peer mediation/conflict resolution, which may have
Evaluation Model 11

positively influenced their growth on the SSBS. Hence, the collection of qualitative data allowed
the evaluators to explain a possible why behind the unexpected outcome at the control school.

_Theoretical Framework._ Another useful part of the design was the theoretical framework.
Aside from providing a structure for thinking about data and a basis for generalizing out
qualitative findings (Seale, 1999), the framework provided a way to organize a summary of the
results of the evaluation. Although the evaluation findings were weighty, this triadic framework
enabled an executive summary format that made for easy reading for the non-educator. Since the
results were disseminated to many non-educators locally, this framework proved useful.

Another benefit to using the context, process, and outcomes model (Silverman, 1993)
was the connections with the conflict resolution process itself, which proved useful to the
schools. One example of this connection is that the framework allowed for conflict resolution to
be recognized as a process that takes time. Caulfield (2000), in her work “Creating Peaceable
Schools,” suggests that “lasting constructive change must be seen as an ongoing process” (p.
184). She further suggests not viewing a peaceable school as a goal, but rather as a process that
takes time. Hall and Hord (2001) concur with the view that change is a process rather than an
event. They indicate in their work on change that focusing on the short-term brings results that
don’t make a significant difference and, therefore, allow evaluators to draw the mistaken
conclusion that the approach does not work. Instead, they suggest that by viewing change at a
school as a process, that change has the time needed to make it be successful. They also indicate
that the time needed is three to five years.

Michael Fullan (2001) further describes the successful change process in his book “The
New Meaning of Educational Change.” He identifies three phases to the change process, each of
which involves time to achieve success. Fullan indicates that the change process is complex and
that it can take three to six years to make changes in schools and up to eight years in a district, depending on the size (Fullan 1999, 2000 as cited in Fullan 2001).

Given the size of the district in which the Allegany Grant was implemented, it was reasonable to assume that the change created by the two-year project was only in the beginning stages described by Fullan. This notion that the process of change takes time—in particular the implementation of conflict resolution to reduce school violence—is one with which the schools have a great deal of difficulty, perhaps due to the political pressure which they are under, albeit the same pressure that causes them to institute the program. Thus, having the theoretical framework in the evaluation model that allows for the recognition of the process of change taking time gives schools evaluation reports framed in a way that can help them buy the time they need to implement change—a useful tool for schools. The results of the two years of evaluations support this need for time, as is seen in the next section.

Macro-level and Micro-level Approach. The macro and micro-level approach was another design strength. It gave both a global and detailed view of the project results. This view was important from a practical perspective. First, the financial cost and time factors involved in completing SSBS assessments on 10,000 students were not realistic. However, completing the assessments on approximately 300 students was manageable.

Furthermore, two years of data documented significant increases in students’ prosocial behavior and significant decreases in students’ antisocial behavior at the classroom level, was supported by qualitative data indicating the development of peaceable classrooms, the first step toward peaceable schools (Lacey & LeBlanc, in press; LeBlanc & Lacey, 2002, LeBlanc & Lacey, 2000). Second, instruments at the macro-level did not indicate any significant changes. Since change in conflict resolution began at the peaceable classroom level, it may be that time
was needed for that change to reach the peaceable school level. Thus, any major evidence of change would not show at the macro-level given the short two-year duration of implementation. Hence, having both a macro or global perspective and a micro or detail perspective enabled a more accurate program evaluation and supported the notion of the time needed for change (Caufield, 2000; Fullan 1999, 2001; Hall & Hord, 2001).

*School Social Behavior Scales.* Finally, using the SSBS as an instrument to rate changes in students’ antisocial behavior and social competence was successful as it allowed for the use of a valid and reliable instrument, pre- and post-test assessment, and treatment and control comparison on the two most important measures of conflict resolution/violence prevention programs. Training teachers how to use the instrument properly was important, as were teacher incentives for timely completion.

*Model Limitations*

*Referral Rate Data.* Referral data as an indicator of program success was not useful. Unfortunately, high mobility rates rendered the referral data unusable. In the ten schools participating in the project over the two-year period, the principal mobility rate was 50% and the assistant principal mobility rate was 30%, which lead to inconsistencies in implementation of the referral system. Teachers’ frequent movement in the schools, ranging from 5% to 31% over the two-year period, similarly impacted the referral rates. Likewise, student mobility, ranging from 31% to 55% of the population over the two-year period, further complicated the referral rate problem, as students may not have been familiar with school rules. These factors also influenced the reporting of referral data by categories. Therefore, the evaluators believe that the referral results reported were not a valid indicator of the success of the conflict resolution program. Given the previous discussion of problems with using referral rates as an indicator of conflict
resolution programs’ success and the compounding problem of mobility rates, the authors do not recommend the use of referrals as part of an evaluation of conflict resolution programs.

**Conflict Resolution Training Pre-test and Post-test.** The conflict resolution training pre-test and post-test was another area of difficulty. They were not administered properly with teacher participants, resulting in unusable data. Furthermore, although the test was to be administered at the parent training sessions, it was impossible to do so.

First, the training had to be trimmed down to an hour-long overview session as that was the maximum amount of time that it was possible to get parents to attend a session. Thus, the test was invalid as it was designed for a full-day training session. Second, the majority of the parents attending the sessions were either illiterate or they did not speak, read, or write English. Given the little amount of time available for training, it was judged inappropriate to take time to assist them with completion of an invalid test.

In the first case, the conflict resolution trainers did not appear to understand the use of the test. In the second case, a unique set of challenges was identified. In both cases the lack of test administration could not have been anticipated as a problem in evaluation model design. However, developing, piloting, and subsequently training trainers on the use of the test need to be part of the evaluation. If parents are illiterate, accommodations should be made (i.e., oral administration and symbols). When parents are literate and the training test is valid, it should be used. Also, the test should be translated into other languages for the population to be served.

**Conclusion**

The major features of the evaluation model are useful in assessing other violence prevention programs. The evaluation model presented in this paper is a large scale, mixed method design using simultaneous methodologies. Having an experimental pre-test and post-test
design with the school as the basis of randomization allows for generalization of the evaluation findings to the larger population. However, due to the realities of schools, unexpected outcomes arise in that a pure control is often not obtainable. Thus, having qualitative data collection enables understanding of the how and why behind evaluation results, particularly unexpected ones. Having a theoretical framework for thinking about data, organizing, and further generalizing evaluation findings to the general population enhanced the design. Collecting data on both the macro and micro levels allowed for global and detailed analysis, while being pragmatic by taking cost and time effectiveness into consideration. Using this model enabled a view of change for what it is—a slow and incremental process that builds over time.
References


Speirs, R. (1994). *Decreasing suspensions in grades nine through twelve through the implementation of a peace curriculum.* (ERIC Document Reproduction Service Number ED378090)


Figure 1. Evaluation Model Data Sources

<table>
<thead>
<tr>
<th>Quantitative Data Sources</th>
<th>Qualitative Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Profiles</td>
<td>Interviews</td>
</tr>
<tr>
<td>- Demographics</td>
<td>- Assistant principals</td>
</tr>
<tr>
<td>- Economics</td>
<td>- Teachers*</td>
</tr>
<tr>
<td>Referrals</td>
<td>Focus Groups</td>
</tr>
<tr>
<td>- School-wide totals</td>
<td>- Teachers*</td>
</tr>
<tr>
<td>- School-wide totals by specified categories</td>
<td>Observations</td>
</tr>
<tr>
<td></td>
<td>- School-walk with field notes</td>
</tr>
<tr>
<td></td>
<td>- Teachers’ classroom instruction with field notes*</td>
</tr>
<tr>
<td>School Climate Surveys</td>
<td>- Florida Performance</td>
</tr>
<tr>
<td></td>
<td>Measurement System*</td>
</tr>
<tr>
<td>Parent Attendance at Training Sessions</td>
<td>- Students’ time-on-task analysis*</td>
</tr>
<tr>
<td>Conflict Resolution Training Pre-test/Post-tests</td>
<td></td>
</tr>
<tr>
<td>School Social Behavior Scales*</td>
<td></td>
</tr>
</tbody>
</table>

Note. * = data sources collected at micro-level, all others at macro-level
Figure 2: Evaluation Model: Strengths and Limitations

<table>
<thead>
<tr>
<th>Model Strengths</th>
<th>Model Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simultaneous Methodologies:</strong></td>
<td><strong>Discipline Referrals:</strong></td>
</tr>
<tr>
<td>+Explains unexplained outcomes</td>
<td>-Impacted by negative school climate</td>
</tr>
<tr>
<td>(i.e., qualitative data explained why significant difference at control school)</td>
<td>-Affected by inconsistent discipline</td>
</tr>
<tr>
<td><strong>Theoretical Framework:</strong></td>
<td>-Influenced by administrators’ styles</td>
</tr>
<tr>
<td>+Provides structure for thinking, organizing and generalizing</td>
<td>-Skewed by mobility</td>
</tr>
<tr>
<td>(i.e., tied conflict resolution to change process and need for time for change)</td>
<td>(i.e., mobility impossible to control for, other factors also can affect results)</td>
</tr>
<tr>
<td><strong>Macro-level and Micro-level Data:</strong></td>
<td><strong>Training Pre-test and Post-test:</strong></td>
</tr>
<tr>
<td>+Provides global and detail views</td>
<td>-Lack of trainers’ understanding</td>
</tr>
<tr>
<td>+Is practical: cost and time effective</td>
<td>-Parental involvement challenges</td>
</tr>
<tr>
<td>(i.e., documented peaceable classroom, peaceable school yet to develop)</td>
<td>(i.e., need for training session with those using test; factor of training duration, need to translate test)</td>
</tr>
<tr>
<td><strong>School Social Behavior Scales:</strong></td>
<td></td>
</tr>
<tr>
<td>+Is valid and reliable</td>
<td></td>
</tr>
<tr>
<td>+Provides pre-test and post-test in one</td>
<td></td>
</tr>
<tr>
<td>+Allows treatment and control options</td>
<td></td>
</tr>
<tr>
<td>(i.e., demonstrated students’ outcomes met; peaceable classrooms attained)</td>
<td></td>
</tr>
</tbody>
</table>