Contents

Introduction .......................................................................................................................... 1
Challenges to Measuring PEEF Casual Impacts................................................................. 1
The Logic Model Approach .............................................................................................. 2
PEEF District-Wide Logic Model ....................................................................................... 3
Recommendations ............................................................................................................. 5
  Communicating PEEF Initiative Results........................................................................ 5
  Recommended Next Steps for Enhancing PEEF Data ..................................................... 12
References .......................................................................................................................... 16
Appendix A. Description of the SRI Matrix of PEEF-Funded Program Measures ........ 19

Figures

1. Public Education Enrichment Fund (PEEF) Initiative Logic Model.......................... 4
2. PEEF Initiative Evaluation Report Suggested Outline............................................... 7


Introduction

The San Francisco Unified School District (SFUSD) contracted with SRI international (SRI) to provide evaluation technical assistance in reviewing current Public Education Enrichment Fund (PEEF) data collection activities. The goal was to help SFUSD refine and streamline current data collection activities so that the data collected can effectively communicate PEEF activities and results. In this report, we provide recommendations that reflect reviews of program materials, previously submitted data, and data collection tools, as well as information gathered through meetings and interviews with PEEF program administrators and managers.

Challenges to Measuring PEEF Causal Impacts

Stakeholders in investment programs such as PEEF frequently want to know, “Did our investments cause improvements in outcomes of participants in the programs we fund?” To make a causal statement, one must rule out all other possible explanations for a change in outcomes, which requires rigorous and often expensive research approaches. The gold standard for measuring causality is a randomized control trial (RCT) in which service recipients are randomly assigned to intervention and nonintervention groups. This approach generates comparison groups that are similar on all other factors except their participation in the intervention being evaluated. Although an RCT is the most scientifically rigorous design for ruling out alternative explanations for impacts observed (Shadish, Cook, & Campbell, 2002), it is expensive, complex, difficult to implement, and most importantly, not feasible for evaluating complex initiatives like PEEF for multiple reasons:

- As an investment portfolio, PEEF has no clearly-defined, separately funded “intervention” to which changes in outcomes can be linked, a requirement of a RCT. PEEF funds mingle with funds from other sources to support a set of widely diverse programs that are intended to work together collectively to improve a wide range of student and staff outcomes district-wide.

- PEEF-funded services often are voluntary and sometimes are mandatory (e.g. physical education), so participants cannot be randomly assigned to receive or not receive PEEF-funded services. The motivations for participants to self-select into a PEEF-funded program make them inherently different from nonparticipants, introducing possible explanations for differences in outcomes other than program participation.

- PEEF explicitly intends to have impacts that spread from individual program participants to others in their schools and district. For example, certain PEEF investments intend to improve overall school climate by providing specific students with peer supports or through training specific teachers in restorative practices. This intended spread of impacts beyond individual participants means that an uncontaminated comparison group whose members are not impacted by any PEEF-funded programs cannot be identified.
A statistical analysis approach that can simulate analyses of data derived from an RCT, propensity scoring, adjusts for differences in program participants and nonparticipants and estimates effects that approximate what might be obtained in an RCT (Cepeda, Boston, Farrar, & Strom, 2003; Rosenbaum & Rubin, 1983, 1985). To use propensity scoring methods, programs need to collect individual, unduplicated data for participants. Currently, only a few PEEF-funded programs (e.g., athletics and the wellness initiative) identify individual program participants. Further, although propensity scoring reduces the bias resulting from self-selection of participants, it may not eliminate it completely, especially when factors such as the spread of effects from program participants to others in a school confounds the differences between participants and nonparticipants. Therefore, propensity scoring also cannot capture overall impact of PEEF investments. Fundamentally, standard program evaluation designs cannot meet the needs of a PEEF evaluation because it is not a program, but an investment portfolio.

Achieving the PEEF evaluation goal of measuring the impact of PEEF as a district-wide investment initiative requires an evaluation design that can measure the synergistic effect within the district of the 14 PEEF-funded programs as a whole. As an appropriate and feasible alternative to creating comparison groups through random assignment or propensity scoring, SRI recommends that PEEF use a logic model approach to assess the results of the overall initiative and the programs its funding helps to support and to examine the relationship of program outcomes to activities and investments.

The Logic Model Approach

The evaluation field promotes a logic model approach for evaluating complex, multifaceted school- or community-wide initiatives because it not only produces useful and credible evaluation results, but also a valuable program or initiative planning tool that is a foundation for monitoring implementation. In fact, PEEF program managers found that the newly-created district-wide logic model helped clarify what the investment initiative intends to accomplish and the model encouraged them to reflect on how the initiative can be strengthened.

A logic model depicts the ways in which various aspects of an initiative connect to one another and how they can be expected to produce the desired outcomes. The recommended approach to constructing a logic model is to first reach agreement with stakeholders as to the outcomes an initiative is expected to achieve and how they can be measured. “Backward mapping” is then used to identify what has to happen—what activities must occur (e.g., a specific program is funded, staff are trained to deliver it, students participate)—if the outcomes can reasonably be expected to occur. Continuing to backup from these “outputs,” the resources, or “inputs” needed to produce the outputs are identified. With these three components identified—resources needed to support activities, which must occur for outcomes to be achieved, and the outcomes desired—stakeholders can monitor both the implementation and outcomes of initiative and their relationship to each other.

Although the logic model approach cannot eliminate all alternative explanations for the outcomes measured in an evaluation, tracking data on each of the logic model components provides credible evidence regarding the actual resources committed, the level of implementation achieved,
and results attained. The logic model approach for establishing impact “draws on tried and true scientific traditions of testing hypotheses about cause and effect relationships, including methods used in physical, biological, and other social sciences” (Connell & Kubisch, 1998, p. 34). Logic models have been used by many other programs and evaluations, including the U.S. Department of Education’s Promise Neighborhood’s effort, Ready Schools Miami, Kellogg Youth Initiative Partnerships, Devolution, ENLACE (Engaging Latino communities for Education), the Native American Higher Education Initiative, and the Build Initiative (Chen & Rossie, 1992; Connell & Kubisch, 1995; Golan et al., 2011; Julian, 1997; Lambur & Mayeske, 2000; Scheirer, 1999; Smith, 2011; W. K. Kellogg Foundation, 2001; Weiss, 1995; Westmoreland, Lopez, & Rosenberg, 2009; Yampolskaya, Nesman, Hernandez, & Koch, 2004). In fact, as part of its effort to improve educational outcomes, the U.S. Department of Education requires successful applicants for many of its grant-making efforts to include a logic model in their proposals and funds the Center for Evaluation and Education Policy to provide technical assistance to applicants in support of logic-model development (http://www.tadnet.org/model_and_performance).

**PEEF District-Wide Logic Model**

Our understanding of key PEEF initiative components and the relationships among them is graphically represented in the PEEF logic model (Figure 1). Because PEEF is a district-wide initiative, this model measures the synergistic effect on the district of the 14 PEEF-funded programs as a whole rather than focusing on each program separately.

The PEEF logic model describes a chain of reasoning or "If… then…” statements that explain how PEEF-funded programs and activities are linked to SFUSD school and student outcomes. There are five components in the model (inputs, outputs, immediate outcomes, and two long-term outcomes), each of which is connected to the next component and is dependent on the component before it occurring.

Beginning at the left of the logic model, the PEEF logic model depicts the 14 PEEF-funded programs (inputs)—the 4 SLAM programs and the 10 Third-Third programs. If these program inputs are funded, then we would expect increases in access and equity to be produced as outputs of what those programs do—their offerings and activities, including for example, the number and types of services, events, and professional development trainings; and the numbers and types of course offerings and equipment provided; and the number and types of program participants (students, parents, and staff).

If outputs are provided and students and staff participate in them, then a set of immediate outcomes would be expected to improve. The four types of immediate outcomes specified in the logic model include: 1) increased professional and instructional capacity and quality (e.g., staff qualifications, skills, and knowledge); 2) improved school climate (e.g., greater safety and cleanliness, reduced behavioral incidents); 3) increased school engagement (e.g., increased...
student attendance, motivation, joy in learning, and leadership skills, and greater parent participation); and 4) improved student health (e.g., self-confidence, self-efficacy, level of stress, fitness and health knowledge and behavior).

*If* progress is made toward these interim outcomes—schools deliver higher quality instruction and have more highly qualified staff, more students feel safe in clean, orderly, and well-maintained schools, more students and parents are engaged and connected to schools, and more students are in good physical and mental health—then it is reasonable to expect that student academic achievement will improve. Extensive research has demonstrated significant links between these four types of immediate outcomes and increased academic achievement. For example, increased teacher qualifications and professional capacity have been related in multiple studies to improved student reading and math achievement levels, higher scores on high school graduation tests, and reduced gaps in student learning frequently associated with differences in household income and race/ethnicity (Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008; Heck, 2007; Powers, 2003; Rivkin, Hanushek, & Kain, 2005; Siegrist, Weeks, Pate, & Monetti, 2009). Aspects of school climate, including students’ feeling safe at school, fewer school-wide behavioral incidents, and overall cleanliness have been associated with increased achievement, including higher grades, lower absenteeism, stronger performance on standardized achievement tests, and high school completion rates (Brand et al., 2008; Brand, Felner, Shim, Seitsinger, & Dumas, 2003; Duran-Narucki, 2008; McEvoy & Welker, 2000; Mendel & Heath, 2004; Samdal, Nutbeam, Wold, & Kannas, 1998). Students’ school engagement, school connectedness, and
motivation toward school also have been demonstrated to have significant impacts on academic achievement, particularly for at-risk urban students (Bond et al., 2007; Fenzel & O'Brennan, 2007; Malecki & Demaray, 2006; Samdal, et al., 1998). Improved student social-emotional and physical health and fitness have been found to be positively related to academic outcomes, including grade performance and high school persistence (Efrat, 2011; Multon, Brown, & Lent, 1991; Murray-Harvey, 2010).

Finally, if increases in academic achievement are attained, then we can expect that students are well on their way to becoming successful adults, ready for college and careers, and equipped with the skills, capacities, and dispositions necessary for 21st century success—SFUSD’s Vision for Student Success.

The combination of “if-then” relationships and measures presented in the PEEF logic model can collectively provide strong and compelling evidence about the success of PEEF. If positive changes in immediate outcomes and academic achievement are documented and these improvements are shown to coincide with increases (or in some cases maintenance in the face of budget cuts from other funding sources) in the number and types of PEEF-funded activities and students and staff served, then it is reasonable to conclude that PEEF played an important role in that impact. The logic model approach provides a standard of evidence that will be convincing.

Recommendations

SRI’s recommendations are presented in two categories. The first set of recommendations applies to what can be done now to communicate PEEF initiative results, based on the current data collection system. The second set of recommendations applies to steps that SFUSD and PEEF administrators may want to consider to further increase their ability to document and communicate PEEF-funded results as the initiative moves forward.

Communicating Current PEEF Initiative Results

We recommend that SFUSD repackage the information in the PEEF performance charts so that it aligns with the PEEF initiative logic model. That is, we suggest that data related to the performance measures (outputs and outcomes) be categorized according to the logic model components, rather than as individual PEEF-funded programs. This approach will more effectively communicate the PEEF initiative’s impacts on students and staff participants. To do this, the PEEF initiative’s report should follow an outline that parallels the causal pathway in the logic model. Figure 2, presents a suggested outline of the proposed report. The PEEF initiative’s report would rely heavily on SRI’s Matrix, the excel database of PEEF-funded program measures, submitted as an addendum to this report. This matrix can be used as a map to help identify the measures related to each area of the logic model. We suggest that each section of the report begin with an evaluation question.
Below, we describe the proposed sections of the report, including suggested evaluation question headings. In addition, we provide information about how to use the SRI matrix of PEEF-funded program measures to identify the measures that should be included in each section of report. The organization of the SRI matrix is included in Appendix A.

**Inputs Section: What does the PEEF initiative provide?**

The report would begin with description of PEEF inputs. This would include brief narrative descriptions of each of the 14 PEEF-funded programs and the overarching goal that each program is trying to meet. The description of each of the 14 PEEF-funded programs should include a brief narrative description of the program’s criteria for allocation of resources.

**Using the SRI Matrix to Identify PEEF-Funded Program Goal Statements.** Draft goal statements for each of the PEEF-funded programs are included in the first row of the SRI Matrix for each program (highlighted in yellow), in the first column—column A, Measure ID. The rows with overarching goal statements are identified by a three-letter program code (e.g. ATH for Athletics) and the number 0 (e.g. ATH-0). The program goal is presented in column B (Goal Text); for example, in row ATH-0, the goal in column B is, “Athletics Overarching Program Goal: Provide equity and access to safe and well-equipped athletic programs at middle and high schools in SFUSD.”

**Outputs Section: What are the levels of access and equity related to the PEEF initiative?**

The next section of the PEEF initiative’s report would communicate the combined impact of PEEF-funded programs in increasing access and equity to sports, libraries, performing arts, and other general education activities across the district by summarizing PEEF-funded outputs. The outputs section of the report would be organized by the four themes included in the PEEF initiative logic model, specifically, increased professional and instructional capacity/quality, improved school climate, increased school engagement, and improved student health. Each of the four sub-sections in the outputs section would begin with a theme-specific output/access/equity question,

1. What are the levels of offerings and participation related to professional and instructional capacity and quality, and their distribution
2. What are the levels of offerings and participation related to school climate, and their distribution?
3. What are the levels of offerings and participation related to school engagement, and their distribution?
4. What are the levels of offerings and participation related to student health, and their distribution?

Performance measures data related to each of these themes should be listed under each question. Program-specific measures should be listed first, then when available, district-wide analyses
Figure 2. PEEF Initiative Evaluation Report Suggested Outline

I. Overview. Goals of the initiative and an explanation of the logic model.

II. Inputs. What does the PEEF initiative provide?
   - Narrative description of 14 PEEF-funded programs, including description of criteria for allocation of resources.

III. Outputs. What are the levels of access and equity related to the PEEF initiative?
   - What are the levels of offerings and participation related to professional and instructional capacity and quality, and their distribution?
     - PEEF-funded program measures data
     - District data
   - What are the levels of offerings and participation related to school climate, and their distribution?
     - PEEF-funded program measures data
     - District data
   - What are the levels of offerings and participation related to school engagement, and their distribution?
     - PEEF-funded program measures data
     - District data
   - What are the levels of offerings and participation related to student health, and their distribution?
     - PEEF-funded program measures data
     - District data

IV. Outcomes: What are the outcomes associated with the PEEF initiative?
   - Immediate Outcomes:
     - Has professional and instructional capacity/quality increased?
       - PEEF-funded program measures data
       - District-wide data
     - Has school climate improved?
       - PEEF-funded program measures data
       - District-wide data
     - Has school engagement increased?
       - PEEF-funded program measures data
       - District-wide data
     - Has student health improved?
       - PEEF-funded program measures data
       - District-wide data
   - Long-term Outcomes:
     - Has academic achievement increased?
       - PEEF-funded program measures data
       - District-wide data

V. Conclusions

Appendix: PEEF initiative budget information
related to an output area (e.g., analyses of the California Health Kids Survey) should be included. As indicated in the four theme-specific questions, PEEF output/access/equity data should include three types of measures, including: offerings and activities (e.g., number of trainings offered, number of classes offered); participation (e.g., number of teachers who attended the trainings, number of students participating in athletics), and distribution (e.g., percent of middle schools with visual and performing arts teachers, percent of translation services provided to Chinese-speaking parents as compared with Laotian-speaking parents. Distribution measures address the issue of equity (e.g. are some schools or some students more likely to get various types of services or more likely to experience certain outcomes than other schools or students?) Most of the analyses related to distribution would be conducted at the PEEF district office. SRI recommendations related to analyses are described on page 13, section c.

**Using the SRI Matrix to Sort for Output Measures by Theme.** Output performance measures are coded as a yes – “Y” in column I, Access/Equity/Output, of the SRI Matrix. Each of the output measures in the SRI matrix also are coded as a yes (Y) in one of the four theme areas: column K, Increased Professional and Instructional Capacity; column L, School Climate; column M, School Engagement; or column N, Student Mental and Physical Health. To identify output measures by theme area, first select the full SRI Matrix excel spread sheet (click in arrow in top left corner of the page to select the whole chart). Then click on the Sort & Filter down arrow, and then click on the Custom Sort option. Enter column I in the “column” box and sort by Z to A in the “order” box, so that the “Yesses” are at the top of the column. Then click on the “add level” box, which will provide the ability to sort by additional column. Use this space to enter the information about the theme area you are describing. For example, for the Increased Professional Capacity section of the output section of the report, enter column K in the “column” box, and enter Z to A in the “order” box, so that again the Yes responses are at the top of the column. Then press OK.

For the Increased Professional Capacity example, all PEEF-funded program measures that have been coded as output measures will be at the top of column I and all measures that have been coded as Increased Professional Capacity measures will be at the top of column K. The Measure ID included in column A, are the measure numbers used by PEEF-funded programs in their performance charts. If SRI suggested an additional measure, the Measure ID for the new measure will include the word “ADD.” The current wording of the measure, as presented in the PEEF-funded program performance charts, is included in column D. If SRI has suggested any changes to this wording, the new wording is included in column F.

The following section provides examples of PEEF-funded program output measures related to offerings and activities and to participation for each of the four theme areas.

1. **What are the levels of offerings and participation related to professional and instructional capacity and quality, and their distribution?**
   - Counts of offerings and activities
   - ATH-2A. Number of PEEF funded professional development trainings provided for coaches
• LS-1C. Average number of hours per week that school libraries are open with a
teacher librarian present
• VPA-1A; Number and percent of schools with credentialed FTE arts teachers who
are PEEF funded (at ES, MS, and HS)
Counts of student, teacher, staff, parent, school participants
• CTE-4Add1; Number and percent of teachers who attend at least one CTE
professional development training
• RP-1A; Number and percent of school site and centralized staff receiving
introductory or advanced training in restorative practices as alternative methods of
discipline

2. **What are the levels of offerings and participation related to school climate, and their
distribution?**
Counts of offerings and activities
• ATH-2F; Number of athletic events covered by private security guards and San
Francisco Police School Resources Officers supported by PEEF funds
• CS-1A; Number and percent of elementary schools receiving an additional .5 FTE
PEEF funded custodian
• PR-2A; Number and percent of SFUSD schools that offer conflict mediation
(overall and by MS, HS)
Counts of student, teacher, staff, parents, school participants
• Currently there are no PEEF-funded program measures of participation in activities
by students, teachers, staff, parents, etc.

3. **What are the levels of offerings and participation related to school engagement, and their
distribution?**
Counts of offerings and activities
• TIS-1A; Number and percent of translation requests from school sites and central
office fulfilled, by language
• PR-1A; Number and percent of SFUSD schools (MS, HS) with a peer mentoring
program
• ATH-2C; Number and percent of bus trips for athletic teams funded by PEEF
Counts of student, teacher, staff, parent, school participants
• SSP-2A; Number of students participating in mentoring programs
• PR-1E; Number and percent of students who are peer mentors/peer leaders (overall
and by MS, HS with a peer mentoring program)

4. **What are the levels of offerings and participation related to student health, and their
distribution?**
Counts of offerings and activities
• SSP-1D Number and percent of students who SSP services referred to student
assistance programs/student success teams
• WI-1A; Number of direct services hours provided to students at the PEEF funded
wellness centers are open
Counts of student, teacher, staff, parent, school participants
**Immediate Outcomes Section: What are the outcomes associated with the PEEF initiative?**

This section of the report should focus on immediate outcomes and also should be organized by the four theme areas. Similar to the output section of the report, data presented in each of the outcome four theme areas should be preceded by an outcome-specific evaluation question.

Performance measures data related to each of these themes should be listed under each question. Program-specific measures should be listed first, then when available, district-wide analyses related to an output area should be included.

**Using the SRI Matrix to Sort for Immediate Outcome Measures by Theme.** Outcome performance measures are coded as a yes – “Y” in column J, Outcome Measure, of the SRI Matrix. Each of the outcome measures in the SRI matrix also are coded as a yes (Y) in one of the four theme areas: column K, Increased Professional and Instructional Capacity; column L, School Climate; column M, School Engagement; or column N, Student Mental and Physical Health. To identify outcome measures by theme area, follow the directions included in the output section, but instead of first sorting measures by output column I, sort measures by outcome column J, then by one of the four theme areas.

The following section provides examples of PEEF-funded program outcome measures for each of the four theme areas.

1. **Has professional and instructional capacity/quality increased since the advent of PEEF?**
   - FA-1B, Number and percent of teachers and principals with at least one log-in to Data Director
   - LS-2C, Number and percent of classroom teachers (overall and by ES, MS, HS), formally collaborating with Teacher Librarians
   - TR-1B, Number and percent of newly hired teachers who are Hispanic (as well as #/% of newly hired teachers who are Hispanic and highly qualified)

2. **Has school climate improved since the advent of PEEF?**
   - ATH-2G, Number of disruptive or violent incidents that were resolved by security and police athletic events.
   - CS-1B, Number and percent of ES principals who are satisfied or very satisfied with overall cleanliness of school site
   - RP-1D, Number and percent SFUSD middle school site and centralized staff who report using aspects of restorative practices
3 Has school engagement increased since the advent of PEEF?
   • ATH-3B, Attendance rates for athletes (overall, and by grade level, gender and ethnicity)
   • PR-2C, Percent of students who, after participating in peer resources, report increased school connectedness
   • TA-2B, Mean attendance of TA aides (those enrolled in TA classes) (overall, and by grade level, gender and ethnicity)

4 Has student health improved since the advent of PEEF?
   • PE-3D, CHKS (ES_Q51) Number and percent of students who report exercising, dancing, or playing sports at least 3 times a week
   • PR-2B, Number and percent of peer mentors/leaders who, after participating in peer resources, report increased ability make a difference in their community (make a difference composite score)
   • WI-2A, Number and percent of wellness clients (students who received 5 or more counseling sessions) who agree of strongly agree that they learned information about how to improve their health and well-being

Long-Term Outcomes Section: Has academic achievement increased?

This section of the report should describe the long-term outcomes related to academic achievement. Program-specific performance measures related to academic achievement should be listed first, then when available, district-wide analyses related to academic achievement (e.g., GPA, test scores) should be included.

Using the SRI Matrix to Sort for Academic Achievement Outcome Measures. Outcome performance measures are coded as a yes – “Y” in column J, Outcome Measure, of the SRI Matrix. Academic achievement measures in the SRI matrix also are coded as a yes (Y) in column O, Academic Achievement. To identify academic achievement outcome measures, follow the directions included in the output section, but instead of first sorting measures by output column I, sort measures by outcome column J, then by academic achievement column O.

Has academic achievement increased?
   • ATH-3A, Mean GPA for athletes (overall, and by grade level, gender and ethnicity)
   • CTE-3B, Mean GPA for students enrolled in career technical education courses (overall, and by grade level, gender and ethnicity)
   • TA-2D, Average number of college units earned by graduating TA aides (overall, and by grade level, gender and ethnicity)

Conclusion Section

This section of the report should summarize overall findings. If positive changes in academic achievement and other desired immediate outcomes are documented and these improvements are shown to coincide with increases in the number and types of PEEF-funded activities and students and staff served, then it is reasonable to conclude that PEEF played an important role in that outcome.
Budget Data Appendix

Budget information for the overall PEEF initiative as well as for each of the 14 PEEF-funded programs should be presented in this section of the report.

Recommended Next Steps for Enhancing PEEF Data

To test the pathway in the PEEF Logic Model, PEEF must collect and report compelling measures of outputs (e.g., the number of service and activity units provided and the number of students and staff participants) in addition to measures of outcomes. Below, we outline recommendations for data and analysis that will allow evaluation of how well the data submitted by the 14 PEEF-funded programs fit the pathway presented in the logic model. In addition to these recommendations, the SRI Matrix appended to this report includes recommendations specific to each of the more than 230 PEEF-funded program measures.

1. Update current measures and processes
   a. Prioritize. Work with program managers to select those measures that best reflect the activities of the program and best convey program goals. With limited resources it is important to target core important questions.
   b. Report unduplicated counts and percentages of total population (e.g., percent of students, or percent of schools, or percent of teachers) wherever possible.
   c. Consider limited and focused qualitative efforts. Quotes from students, teachers, principals and parents on how programs meet their needs can be included in reports to enhance and provide depth and detail to the quantitative data. Some of this data may be available already, e.g., prior interviews conducted by the CAC.

2. District-level data analyses issues/options
   a. Presenting data related to change (increases/decreases) in outputs or outcomes requires longitudinal or trend data for comparisons across time (e.g. prior to PEEF funding vs. after, or tracking yearly data since implementation). If you don’t currently have prior year data for a measure, consider asking program managers to go back in time if possible, to update/recalculate data to match current measurement recommendations.
   b. Consider keeping the metric consistent when making comparisons across time, in light of decreasing PEEF funding. To control for PEEF funding decreasing over time consider analyzing the number of students who participated per $1 million (or per $1 thousand dollars) of funding, or the number of classes offered per million (or thousand) dollars of funding. For example, if you served 2500 students in 2008 for $52,000, divide 2500 students by $52,000 = .048 students/thousand dollars. If in 2011 you served 2,200 students for $40 thousand, then divide 2,200 students by $40,000 = .05 students/thousand dollars. Even though you served fewer students in 2011 (2,200 students) than in 2008 (2,500 students), you actually served more students per thousand dollars spent (.048 students/thousand vs. .050 students/thousand).
c. Distribution measures address the issue of equity (e.g. are some schools or some students more likely to get various types of services or more likely to experience certain outcomes than other schools or students?) To address these types of questions, when individual-level data (with student HO numbers) is available from programs, consider analyzing and presenting the data by various subcategories, including,
   i. gender,
   ii. race/ethnicity,
   iii. ELL status,
   iv. school type (e.g. elementary, middle, high school),
   v. Zone, etc.
In addition, when reporting school-level data, consider reporting measures by school type and/or by school Zone, particularly when addressing distribution of PEEF-funded offerings and activities (for example comparison of percent of elementary, middle, and high schools with FTE arts teachers or Zone to Zone comparisons, comparing data across superintendent zones).

d. Analyses that demonstrate the proportion of SFUSD schools, students, and/or staff overall who have been touched by at least one PEEF program would provide a strong measure of the initiative’s impact.

e. Present district-wide academic data by school type, Zone, etc.:
   i. GPAs
   ii. Attendance data,
   iii. test scores,
   iv. graduation rates/ dropout rates,
   v. suspension and expulsion rates.

f. Analyze extant data in relation to PEEF outputs (several suggested items from these surveys have been added as measures to the SRI matrix).
   i. Possible sources of extant data
      1. California Health Kids Survey,
      2. Youth Risk Behavior Survey
      3. Principal Survey.

g. Proposed analysis comparison options.
   i. Use time series analysis (longitudinal analysis) wherever possible, showing trends before and annually after the advent of PEEF.
   ii. School to school comparisons (Comparisons can be made by examining changes over time for specific sets of schools and also by comparing the changes overtime for schools that received certain services/levels of funding with schools that did not receive those same services/funding):
      1. schools receiving PEEF funds (for a particular activity) vs. schools that are not receiving PEEF funds (for that activity),
      2. schools that receive more or less PEEF funds per student capita.
   iii. Student to student comparisons
      1. students participating in PEEF vs. students who are not participating in PEEF funded offerings/activities.
iv. Within individual student comparisons – before and after involvement with PEEF (e.g., comparing GPA of individual student before and after his/her involvement as a Teacher Academy aide). This may be particularly useful in situations where you wouldn’t expect the GPAs or attendance rates of PEEF participants to be different/better than other students, overall.

3. Future Considerations
   a. Ideally, all programs should collect both output and outcome data. Currently some of the programs include only one type of measure.
   b. Consider collecting a small but consistent set of core measures across programs on students and staff served. Core measures should include outputs as well as key outcomes that PEEF intends to affect overall. Base the selection of these measures and the type of data you collect on the logic model so that you can adequately document each step in the causal pathway. Some items/measures to consider collecting include students name and/or HO number so that it can be linked to other district databases, information on dosage of services received, and key outcome questions that can be asked of all participating students.
   c. Consider collecting HO numbers for students whenever possible.
   d. Build capacity of program managers to collect high quality data.
      i. Establish protocols with managers so that data are tracked, stored, and reported in a consistent manner across time and programs.
      ii. Increase district-level funding for PEEF initiative evaluation activities, particularly to support program managers, many of whom indicated during the phone calls their strong need for additional support in designing their evaluation activities, developing of surveys and data collection tools, sampling, and analyzing data. Data in the PEEF reports are only as strong as the data from the individual programs.
      iii. Consider requiring PEEF-funded programs set aside resources within their program budgets for evaluation activities. Government grants and foundations usually suggest 10-15% of program budgets be set aside for evaluation activities.
   e. Invest in the development of a PEEF data management system that will allow consistent data tracking across programs, as well as the unduplication of counts. It is not sufficient for programs to enter data into performance charts without original data being accessible to PEEF staff. Make sure that development of the database does not duplicate work of some programs that have databases (like Wellness). Find way to mesh the systems together.
   f. Redesign existing surveys.
      i. Redesign district-wide PEEF Principal Survey (consider survey monkey?) (issues include being able to disaggregate results so can see results of PEEF funding/relative PEEF investment). Measure principal perceptions of outcomes experienced since the advent of PEEF.
ii. Consider revising the one PEEF question on the district wide student survey (and lobby to add a few more).

  g. Data should be available at least annually (e.g., conduct Wellness surveys yearly, not biennially).

  h. Consider evaluating select PEEF programs using propensity score matching techniques to create a student comparison group.

  i. Explore further the option of comparing district trends to trends in another comparable school district that did not receive PEEF-type funds. Likely will be very difficult to find appropriate comparisons.

In the matrix, we suggest "present as a cumulative total," for various measures where the good/service provided is a one-time provision. For example, ATH-1F.# of Ath facilities improved by PEEF, or LS1-D (# and percent of libraries that have received a technology upgrade).
References


Cepeda, M. S., Boston, R., Farrar, J. T., & Strom, B. L. (2003). Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *American Journal of Epidemiology, 158*, 280-287.


Appendix A

Description of the SRI Matrix of PEEF-Funded Program Measures
Organization of the SRI Matrix

The SRI Matrix is organized in an Excel spreadsheet. We included every measure from every program in the matrix, by row. Each measure is listed in its own row. We listed them in the order they are listed in each program’s performance chart. Before a program’s measures are listed we include a row where we list a draft of the program’s overarching goal. This is our understanding of each program’s “mission,” derived from reviewing the programs’ goals and measures, and from speaking with program managers. The rows with the overarching goals are highlighted in yellow.

As described, each measure is listed in its own row. Information related to, or associated with the measure is then presented by column, across the row. Below is the description of each column.

**Column A:** Contains the measure number. These measure numbers are taken directly from the programs’ performance charts (e.g., 1A). SRI used an acronym for each program (e.g., Physical Education: “PE,” and then linked the acronym to all measures for that program (PE-1A). It is important to note that SRI assigned all overarching goals a value of “0,” so PE’s overarching goal would be listed as PE-0. In addition, where SRI suggested adding measures, the measure would be listed by the # of the program measure it is most closely linked to, with the word “Add,” and a number to indicate whether it’s the first measure we’ve suggested adding, the second measure, etc. For example, PE-2BAdd1, PE-2BAdd2, etc.

**Column B:** Contains the goal text of all goals listed on the programs’ performance charts. These goals are linked to their corresponding measures as indicated by the performance charts. Because all goals had multiple measures, the text is repeated in column B for every measure that it is linked to and therefore is repeated for multiple rows.

**Column C:** Contains the goal-related evaluation question. This is the question that SRI proposed as the one that best matches the listed goal. When conducting an evaluation, it is helpful to know what is being asked. When a goal such as: “Improve student attendance, achievement and pro-social behavior,” is given, a question that asks whether this goal has been met might be: “Has SSP improved student attendance, achievement, and pro-social behavior?” SRI proposed a question for each program goal.

**Column D:** Contains the measure text, exactly as it was listed on the program’s performance charts. If the word “Add” appears, then SRI has proposed a new or additional program measures (the text for these additional measures will be listed in column F, as described below, rather than in column D).

**Column E:** Contains the measure-related evaluation question. This is the question that SRI proposed as the way of asking, or understanding, what is being measured. These questions are useful in guiding what data to collect, and how to analyze the data, as it’s necessary to know what question is being asked before knowing how to proceed with answering it.
**Column F:** Contains measure recommendations by SRI. This column contains any proposed revisions to the measure text in Column D, or the text of any additional (new) measures that SRI is proposing.

**Column G:** Contains comments and questions posed by SRI, usually related to the measure.

**Column H:** Contains analysis suggestions by SRI. SRI made suggestions for analyzing most of the measures (e.g., trend analysis of data pre-post PEEF).

**Columns I & J:** Column I is labeled “Access and Equity (Output),” and column J is labeled “Outcome.” Each measure is either an access/equity/output measure or an outcome measure, but not both. Each measure will therefore be coded accordingly, and will contain a “Y” in the appropriate category. For example, if a measure is an output (number of trainings provided), there will be a “Y” in the Access & Equity (Output) column, and an “N” in the outcomes column. This system was used in order to allow the PEEF district office to sort by outcome and output.

**Columns K-O:** Columns K-O each contain one theme that aligns with the themes listed in the PEEF logic model. Column K is “Increased Professional and Instructional Capacity;” column L is “School Climate;” column M is “School Engagement;” column N is “Student Mental and Physical Health;” and column O is “Academic Achievement.” Each measure is coded “Y” for the theme it most closely aligns with. It is important to note that each measure is only coded “Y” for one category, and there will only be one “Y,” with all other columns showing an “N.” For example, a measure that looks at number of “professional trainings” would be coded “Y” under “Increased Professional and Instructional Capacity.” This system is intended to allow the PEEF office to sort by logic model theme.