

Program Evaluation: Where Instruction and School Business Meet

A strategy commonly used by instructional leaders can help school business officials improve their budget process.

By Steven V. Ayers, Ed.D., CPA

As districts struggle to develop budgets in these challenging economic times, school business officials might consider turning to program evaluation for help. Program evaluation gives school business officials an opportunity to improve the budgeting process by examining the cost-effectiveness of the district's programs and initiatives.

For an excellent introduction to this emerging field, see Carol Weiss's *Evaluation* (1998). For purposes of this overview, I will discuss four key elements I used to conduct a quantitative evaluation of a class size reduction initiative:

- Literature review
- Logic model/program theory

- Data identification, compilation, and testing
- Evaluation of results and conclusion.

Literature Review

As tempting as it is to dive into the data, it is important to first understand the research that gave rise to the initiative. This research will provide insight into the design of the initiative and its expected outcomes. It will also provide a basis for comparing the manner in which the initiative was implemented with that described in the research. In some cases, lack of fidelity to the original research may inhibit the replication of the research results.

It is also useful to look at contradictory research results to determine any criticisms of the initiative and incorporate them into the testing design. Google Scholar is an excellent search site for relevant research articles.

Performing the literature review enabled me to move beyond my cursory level of knowledge about the effects of reduced elementary class sizes to a detailed understanding of the major studies that gave rise to its popularity, the relatively small effect with which size was associated, and implementation problems educators encountered in other states. My review also led me to identify differences in the manner in which reduced class size was implemented in my district versus its most successful applications, perhaps explaining why the results that I observed locally were not as robust or sustained as those noted in the literature.



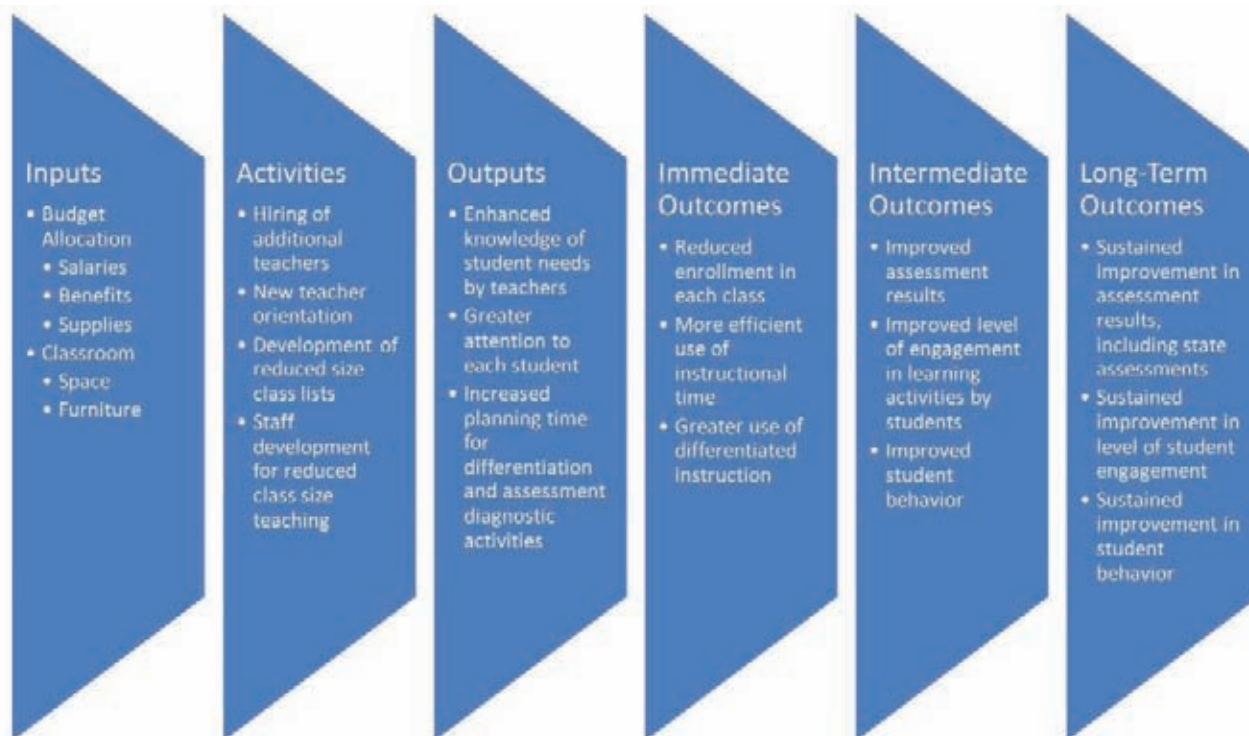


Figure 1. Logic model addressing reduced class size.

Logic Model/Program Theory

A key element of program evaluation is a well-designed logic model. If you’ve spent time in public accounting, you will remember the flowcharts of client internal control systems you did in your early years. A logic model is a visual display of the various inputs and outputs of an initiative that are believed to give rise to the expected outcome. Figure 1 displays the logic model I used to examine the efficacy of a reduced class size initiative.

Developing a logic model ensures that the evaluator fully understands all the assumptions inherent in the design of the program. It provides the opportunity to discover logic gaps, uncover unrealistic assumptions, and design tests to determine if the model has been implemented as planned or is having the anticipated effects.

A logic model ensures that the evaluator fully understands all the assumptions inherent in the design of the program.

The Kellogg Foundation’s *Logic Model Development Guide* (2004) is an excellent resource for developing and using logic models.

In an ideal world, a program evaluator has a fully developed logic model as part of the implementation plan. Regrettably, this step is rarely a component of the plan.

Too often, schools adopt initiatives with vague goals, such as “improve student achievement,” without any definition of success for meeting that goal. Thus, if one student is performing better after the implementation of the initiative, it could be asserted that the goal was met.

Too often, schools adopt initiatives with vague goals.

Incorporating logic models into the implementation planning identifies these kinds of weaknesses by requiring that the perceived cause-and-effect relationships be clearly illustrated, with identified outcomes, and that the resources and steps that will generate these outcomes are spelled out in adequate detail.

Data Identification, Compilation, and Testing

After the evaluator thoroughly understands the initiative’s theory and anticipated outcome, it is time to develop a testing plan. This task involves determining which data are available to measure effects and which data must be developed.

Data can take many forms and come from many sources, including student test scores, demographics, and attitudes gathered from focus groups and surveys. The data will often be drawn from multiple sources and must be compiled into a single consolidated database.

Microsoft Access or Excel can help accomplish this task. Great care must be taken to preserve the integrity of the data during this phase of the testing.

After the evaluator identifies and compiles adequate measures, the next step is to select appropriate statistical tools to discern whether any observed differences in the data have significance, meaning these differences are unlikely to be attributable to normal population variability. Many programs are available to assist in this exercise. (SPSS is most commonly used in social science settings.) In addition, Microsoft Excel offers many of the most common statistical functions.

The evaluator must always look at an incomplete picture and manage the risk of drawing an incorrect conclusion.

Among the most common tests are *t*-tests for comparing means, ANOVAs (analyses of variance) for comparing multiple means, and inferential statistics, such as correlation and regression, for identifying relationships between inputs and outputs.

The key to an effective analysis of the data is to ensure that the statistical test selected matches the type of data gathered. For example, the tests just described are all most appropriate for interval data, where differences between datum points are fixed and objective. Nominal data, such as a Likert scale, have differences between datum points that will vary by respondent. For such data, a chi-square test is often the most appropriate measurement. For additional guidance on the effective use of data in school districts, see *Schools and Data* (Creighton 2007).

Evaluation of Results and Conclusion

After the data are compiled and tested, the evaluator is prepared to draw a conclusion. In doing so, it is important to understand the limitations on what can be inferred through statistical analysis.

The evaluator is always trying to draw conclusions about the population tested without having the luxury of examining the entire population. Thus, a number of pitfalls must be recognized, including the possibility of researcher bias, a question about whether the data examined actually measure the populations, and confusing correlation and causation. Therefore, the evaluator must always look at an incomplete picture and manage the risk of drawing an incorrect conclusion.

Despite these risks and limitations, the result of this process is a greatly enhanced understanding of the ways in which the initiative affected the organization and whether expectations were met. This understanding may lead to a conclusion that the program was successful and should be retained, or that it did not meet its objective and should be terminated.

The field of program evaluation is in its infancy, particularly in the public education arena.

More often, the results will not be as clear, but will lead to other areas for exploration, both from a formative perspective focused on the manner in which the program was implemented and from additional substantive testing with refined measurements based on what was learned in the additional evaluation process. Regardless of the final conclusion, the district will benefit from funding decisions that are based on objective evidence, thus leading to continued organizational improvement.

Final Thoughts

The field of program evaluation is in its infancy, particularly in the public education arena. The continued reduction in state and local funding makes it imperative that districts make sound data-driven funding decisions. The abundance of data as a result of No Child Left Behind makes this type of decision making more feasible than ever.

The barrier in all but the largest districts is the shortage of people with the necessary skills to perform this analysis. School business officials have these skills and have an opportunity to take on a new and vitally important role in education leadership.

References

- Creighton, T. B. 2007. *Schools and data*. Thousand Oaks, CA: Corwin Press.
- W. K. Kellogg Foundation. 2004. *Logic model development guide*. Battle Creek, MI: W. K. Kellogg Foundation. http://ww2.wkcf.org/DesktopModules/WKF.00_DmaSupport/ViewDoc.aspx?LanguageID=0&CID=284&ListID=28&ItemID=2813669&file=PDFFile.
- Weiss, C. 1998. *Evaluation*. Upper Saddle River, NJ: Prentice Hall.

Steven V. Ayers, Ed.D, CPA, is assistant superintendent for business for the Hilton Central School District in New York. He is also school business administration program coordinator for State University of New York College at Brockport. Email: sayers@hilton.k12.ny.us

COMPUTERS IN THE SCHOOLS

In fall 2008, all public schools in the United States had one or more instructional computers with Internet access, according to the National Center for Education Statistics. The ratio of students to instructional computers with Internet access was 3.1 to 1. Ninety-seven percent of schools had one or more instructional computers located in classrooms (excluding laptops on carts) and 58% of schools had laptops on carts.