Appendix LLL-12-18
June 14, 2010

MMSD Program Evaluation Protocol

Report to the Board of Education – June 14, 2010

I. Introduction

A. Title or topic - District Evaluation Protocol – The presentation is in response to the need to provide timely and prioritized information to the Board of Education around programs and interventions used within the District. The report describes a recommended approach to formalizing the program evaluation process within the District.

B. Presenters

Kurt Kiefer – Chief Information Office/Director of Research and Evaluation
Lisa Wachtel – Executive Director of Teaching & Learning
Steve Hartley – Chief of Staff

C. Background information - As part of the strategic plan it was determined that priority must be given to systematically collect data around programs and services provided within the district. The purposes for such information vary from determining program and intervention effectiveness for specific student outcomes, to customer satisfaction, to cost effectiveness analyses. In addition, at the December 2009 Board meeting the issue of conducting program evaluation in specific curricular areas was discussed. This report provides specific recommendations on how to coordinate such investigations and studies.

D. Action requested - The administration is requesting that the Board approve this protocol such that it becomes the model by which priority is established for conducting curricular, program, and intervention evaluations into the future.

II. Summary of Current Information

A. Synthesis of the topic - School districts are expected to continuously improve student achievement and ensure the effective use of resources. Evaluation is the means by which school systems determine the degree to which schools, programs, departments, and staff meet their goals as defined by their roles and responsibilities. It involves the collection of data that is then transformed into useful results to inform decisions. In particular, program evaluation is commonly defined as the systematic assessment of the operation and/or outcomes of a program, compared to a set of explicit or implicit standards as a means of contributing to the improvement of the program.

Program evaluation is a process. The first step to evaluating a program is to have a clear understanding of why the evaluation is being conducted in the first place. Focusing the evaluation helps an evaluator identify the most crucial questions and how those questions can be realistically answered given the context of the program and resources available. With a firm understanding of programs and/or activities that might be evaluated, evaluators consider who is affected by the program (stakeholders) and who might receive and or use information resulting from the evaluation (audiences). It is critical that the administration work with the

Board of Education on clearly defining the key questions any evaluation is designed to answer to assure that what is produced meets expectations.

Whether the evaluation is being conducted in order to determine success or failure (summative evaluation) of a program, or to make improvements through adjustments based on ongoing feedback (formative evaluation), planning the evaluation includes developing processes to understand the target audience, developing meaningful program objectives, and selecting appropriate indicators to answer questions. An effective evaluation should identify if the program has been implemented as intended and has produced desired outcomes. As prioritizing evaluations can be challenging for a school district with many programs, there are several considerations that may be weighed when determining stakes of programs and their outcomes including:

- **Program cost** – Programs that are expensive need to be proven effective and if not improved or abandoned.
- **Importance of outcomes** (e.g., implications of program failure) – Certain programs have serious implications for failure.
- **Perceived importance of program/outcomes by stakeholders and audiences** – In some cases the reason a program is being evaluated has to do with a request by an audience (e.g., a funding source).

B. Recommendations – The following steps are being recommended to formalize the MMSD evaluation protocol. The recommendations were informed in large part by the work commissioned to Hanover Research Council. The HRC study included contacting several K-12 districts across the country to determine current and best practices.

1. Curricular Program Review Cycle - A key part of the overall district evaluation strategy must include a regular curricular program review. Curricular areas recommended for review include literacy, math, science, social studies, world languages, the arts, health and physical education, and career and technical education. Each curricular area would rotate through a cycle of review on a seven year basis. The stages of the review include:

    - **Year 1** – Evaluation design and preliminary data collection, evaluation committee established and oversight tasks, Evaluation Year/Data interpretation, report and recommendations
    - **Years 2/3** - Refinement of evaluation design and data collection based on continuous feedback and oversight, review and select curriculum resources, conduct professional development
    - **Year 4** – Program revisions and implementation of curriculum, additional professional development, on-going monitoring
    - **Year 5** – Additional professional development and on-going monitoring
    - **Year 6** – Continued professional development, preparation for year 1 program evaluation cycle tasks

2. New Programs and Interventions - All new programs and interventions should be selected based on the existing evidence of success. Secondary research should be conducted to determine the level of rigor of existing evidence. Quantitative effects on student achievement using randomized trials or quasi-experimental designs that involve treatment and control/comparison groups should be available. The research should be conducted by non-involved third parties, i.e., not vendors researching their own programs and

interventions. The evidence should also examine whether or not the program or intervention has clearly defined methods for assuring high quality implementation, i.e., fidelity.

When sufficient evidence does not exist regarding a program or intervention outcome son student learning then a study should be completed. The methods of those studies may vary, but quantitative outcomes on student learning should be the goal. Other key questions might also be explored including the quality or fidelity of implementation which are factors that may alter the student learning outcomes. The level of effort and resource commitment would be defined using the protocol established within this document.

3. Evaluation Design Approval - Before commencing a program evaluation, it is recommended that a research design be submitted to the Board of Education for feedback. Instructional committees would serve as advisors the evaluation process. This should take place the year prior to the data collection year to ensure that the evaluation is asking the right questions and so that adjustments can be made as needed before implementation. Each Board of Education sanctioned evaluation design will consist of the following information which shall be reviewed by the Board of Education in the approval process.

   - Define purpose, scope, goals and objectives;
   - Specify the evaluation questions and the limits of the evaluation;
   - Determine the data collection plan including how it will be collected, when, and by whom; and
   - Clarify how the data will be analyzed and outline the elements of the report that will be produced.

   Annually, the Board of Education will review a list of proposed program and intervention evaluations that will be conducted in support of the overall curricular review cycle process or in conjunction with other priorities as they deem appropriate.

4. Collaboration on Conducting Evaluations - To carry out a program evaluation, the Research & Evaluation Office would work in collaboration with the Department of Curriculum and Assessment on a research design which is guided by the evaluation protocol.

5. Reporting - Updates to the Board would be a routine expectation for these evaluation projects while in process as well as in the form of a final report at the conclusion of each evaluation study.

III. Implications

A. **Budget** - Multiple funding sources will be targeted as appropriate to support the various components of the evaluation process and the cycle of curricular content areas. Primary funding responsibility for specific functions will be shared between the Departments of Research & Evaluation and Curriculum and Assessment. In addition, specific funds have been dedicated in the Strategic Plan – Resource and Capacity, Rigorous Evaluation, Action Step #1.

B. **Strategic Plan** – The MMSD Program Evaluation Plans addresses action steps in MMSD's Strategic Plan as described below:

Resource and Capacity, Rigorous Evaluation, Action Step #1: Identify appropriate quantitative and qualitative evaluation methods to answer questions related to the key district goals.

C. **Equity Plan –** The MMSD Program Evaluation Plan addresses the Equity Plan as described below in the Statement of Commitment, page 5:

This commitment to a culture of equity will permeate all interactions, decisions, and practices throughout the District as reflected in, but not limited to, the following measures:

1. **LEADERSHIP.** Active modeling and implementation of equity-minded policies and practices in all aspects of teaching and learning.
2. **ACCESS & ACHIEVEMENT.** Elimination of gaps in access and achievement due to current or historic inequalities.
3. **QUALITY.** Provision of necessary resources and services to meet the needs of all students.
4. **CLIMATE.** Continuous engagement of classrooms, schools, and community.
5. **ACCOUNTABILITY.** Assessment and evaluation of policies and practices demonstrating continuous improvement.

D. **Implications for other staffing**

1. Internal Resources - While the primary tasks of program evaluation lie within the Research & Evaluation Office, partnerships are required with other Departments and Divisions in designing and implementing various evaluation tasks. For example, the curricular review process will rely heavily on the involvement of the Curriculum and Assessment Department. It is anticipated that in order to design and implement up to three specific program and intervention evaluations annually at a minimum one full time equivalent employee is needed within the Research & Evaluation Office. This would include tasks associated data collection, analysis, and reporting. An equal proportion of staff time would be needed in the Curriculum & Assessment Department for the tasks associated with the curricular review process.

2. External Resources - Other options for conducting evaluation by third parties must also be pursued. These include, but are not limited to, the following entities/agencies/organizations:

   a. *MMSD External Research Committee* - This long-standing group reviews and approves all requests by third parties to conduct research within our schools. We suggest more formally posting the District's research agenda topics on its web site. We could also provide an incentive for researchers wishing to conduct research on these topics. We would communicate this research agenda with UW and WCER on a regular basis and ask that they in turn distribute the agenda to stakeholders within the UW.

   b. *Minority Student Achievement Network (MSAN)* – Housed within the Wisconsin Center for Education Research (WCER), MSAN conducts coordinated research across member districts which is guided by a Research Practitioners Council. MSAN regularly creates relationship with third party research entities to carry out it's research agenda. Projects to date have included a study on the affects of homework on mathematics knowledge and skill development.

   c. *Midwest Urban Research Network (MURN)* - Collaborate with Learning Point Associates (formerly Midwest regional Education Lab) on research and evaluation projects. The

collaborative includes similar sized school districts across the Midwest. Their initial project is an analysis of predictive characteristics of students at risk of drop out prior to graduation.

d. *Wisconsin K-12 Program Evaluation Consortium* - This idea has grown out of conversations with WCER. A request has been submitted by WCER in conjunction with MMSD to the W. T. Grant Foundation to investigate the feasibility of creating an entity that involves Wisconsin K-12 districts in conducting program evaluations. Given the limited resources of any single school district in the state, and with the development of a state wide data base of longitudinal student data, it may be feasible to engage in collaborative evaluations into the effects of various interventions.

e. *StateWide Longitidinal Data System(LDS)* – As part of a federal grant obtained by the Wisconsin Department of Public Instruction (DPI), a state wide data warehouse has been created. That data could be used for collaborative research and evaluation projects across the districts in Wisconsin. MMSD has been a strong advocate for greater use of the LDS for these purposes and will continue to push for more projects.

IV. Supporting documentation

Attached appendices include:

Appendix A - Report provided by Hanover Research Council (HRC)
This report was in response to the administration's request in fulfillment of the Board of Education's December 2009 directive. The report summarizes inquiries made of several school districts around the country relative to their program evaluation activities and practices.

Appendix B – Addendum to Report provided by Hanover Research Council (HRC)
This report provided additional information from another K-12 school district which supplied their responses to the HRC survey after the initial report was submitted to MMSD.

Appendix C - Draft MMSD Curricular Review and Renewal Cycle
This document describes the funding, process changes and proposed curricular content areas that will be reviewed over the next six years. This document serves as a transition from primarily a school-based to a district-wide review process.

Appendix D - Summary for Program Evaluation
This report was shared with the MMSD Instructional Council in November 2008. This also serves as context for evaluation tasks and practices within the MMSD.

HANOVER RESEARCH COUNCIL

1101 Pennsylvania Ave. NW, Suite 600 Washington, DC 20004
P 202.756.2971 F 866.808.6585 www.hanoverresearch.com

Appendix A - Report provided by Hanover Research Council (HRC)

District Program Evaluation Practices

Prepared for Madison Metropolitan School District

This report by The Hanover Research Council surveys the program evaluation practices of public schools. Background information on program evaluation processes via materials from university and other education-related organizations are supplemented by examples of districts' methodologies obtained through interviews with relevant school administrators.

Executive Summary

All education programs need to include an evaluation component if their success is to be determined, and if weaknesses in the programs are to be identified and corrected. In the words of the International Institute for Educational Planning, established in Paris in 1963 by UNESCO:[1]

> Education programs cannot be said to be effective if there are no measurable improvements in student learning over time. Similarly, some comparison group, or groups, of teachers and students should be included in the study. If there are measurable changes in student learning over time, but the magnitude of the changes is not different from changes that occur in non-program students, then the program cannot be said to be effective.

School districts are expected to continuously improve student achievement and ensure the effective use of resources. Evaluation is the means by which school systems determine the degree to which schools, programs, departments, and staff meet their goals as defined by their roles and responsibilities. It involves the collection of data that is then transformed into useful results to inform decisions. In particular, program evaluation is commonly defined as the systematic assessment of the operation and/or outcomes of a program, compared to a set of explicit or implicit standards as a means of contributing to the improvement of the program.[2]

Program evaluation is a process. The first step to evaluating a program is to have a clear understanding of why the evaluation is being conducted in the first place. Focusing the evaluation helps an evaluator identify the most crucial questions and how those questions can be realistically answered given the context of the program and resources available. With a firm understanding of programs and/or activities that might be evaluated, evaluators consider who is affected by the program (stakeholders) and who might receive and or use information resulting from the evaluation (audiences).

Whether the evaluation is being conducted in order to determine success or failure (summative evaluation) of a program, or to make improvements through adjustments based on ongoing feedback (formative evaluation), planning the evaluation includes developing processes to understand the target audience, developing meaningful program objectives, and selecting appropriate indicators to answer questions. An

[1] Anderson, L.W., and Postlethwaite, T.N. 2007. "Program evaluation: Large-scale and small-scale studies." International Institute for Educational Planning (UNESCO) Education Policy Series, Vol. 8, p. ii. http://www.iiep.unesco.org/fileadmin/user_upload/Info_Services_Publications/pdf/2007/Edpol8.pdf
[2] Shackman, G. 2010. "What is program evaluation?" The Global Social Change Research Project. http://gsociology.icaap.org/methods/evaluationbeginnersguide.pdf

effective evaluation should identify if the program has been implemented as intended and has produced desired outcomes. As prioritizing evaluations can be challenging for a school district with many programs, there are several considerations that may be weighed when determining stakes of programs and their outcomes including:[3]

- **Program cost** – Programs that are expensive need to be proven effective and if not improved or abandoned.
- **Importance of outcomes** (e.g., implications of program failure) – Certain programs have serious implications for failure.
- **Perceived importance of program/outcomes by stakeholders and audiences** – In some cases the reason a program is being evaluated has to do with a request by an audience (e.g., a funding source).

A total of seven districts were surveyed by Hanover to get a better understanding of how they prioritize and conduct program evaluations. We targeted 17 large urban and suburban districts across the country with prominent research/evaluation departments. Our goal was to gather the input of the directors or other key personnel of these departments on the types of evaluations conducted, the people involved, and the evaluation process used. Based on our interviews and a review of materials posted on the districts' websites, we find on the whole that:

- Research/Evaluation departments have multiple responsibilities, including (but not limited to) data procurement, management, and interpretation; research consultation to schools; coordination of outside research requests; and in some cases administering district-wide tests.
- These departments are typically small among surveyed districts but are staffed with individuals highly trained in research methods.
- Much of the departments' resources are spent complying with data requests of district leaders and teachers rather than on formal program evaluations.
- Data warehouses serve as a foundation for the districts' evaluations and other accountability assessments.
- Satisfaction surveys are a common tool employed by districts to assess stakeholders' views on programs.
- Program evaluations often follow a district's textbook adoption cycle, which ranges from five to six years for surveyed districts.
- External evaluators are contracted more frequently for programs that are funded by grants, though it is often a collaborative process with internal staff.

[3] Aldrich, S. n.d. "Program evaluation planning and design: A Guide for teacher centers" New York State Teacher Resource Centers, p. 3. http://www.programevaluation.org/docs/PEplantut.pdf

- True experimental designs for program evaluations are not the norm; exploratory and quasi-experimental research designs are more common due to limited time and resources.
- Timelines for carrying out program evaluations are not standard but more so based on school leaders' needs for information.
- Funding for program evaluating is a fraction of departments' budgets—which are usually small to begin with—and the districts do not budget for specific evaluations but instead draw from resources as needed.

Given these observations, this report will provide further insight into the program evaluation practices of public school districts.

Section One examines the recommended steps to evaluating a program through a review of literature covering the importance of data for school improvement, evaluation planning, and a general framework for conducting evaluations.

Section Two explores some examples of implemented evaluation strategies of district programs. As noted above, we gathered the examples from publicly available sources and interviewed school leaders for information about their district's evaluation process.

Section One: Program Evaluation Steps for School Improvement

In an age of accountability, data is seen as the driving force behind school program improvement. It is used by districts to plot progress, plan and execute instructional interventions, report results, and hold students, teachers, administrators, and school systems accountable. Meaningful data collection and analysis helps districts make decisions about policies, programs, and individual students. This section discusses the role that data play in districts' decision making about their school programs, considerations by districts before initiating program evaluations, and a general framework for conducting evaluations.

Importance of Data to Evaluations

Data-driven decision making (DDDM) is a system of teaching and management practices that gets better information about students into the hands of classroom teachers.[4] The RAND Corporation, a California-based non-profit research organization, defines DDDM in education as the systematic collection and analysis by teachers, principals, and administrators of various types of data - including input, process, outcome and satisfaction data - to guide a range of decisions to help improve the success of students and schools.[5]

The impact of school programs is based on data. When a district wants to know how literacy, science, or other programs being used in its schools are affecting students' learning, it turns to its staff or outside firms to carry out an evaluation of the program. Evaluations can provide useful information about what is happening in the school and a strong, data-driven foundation for designing, implementing, and improving strategies that promote student achievement.[6] Concerning school programs, evaluations are useful to:[7]

- Improve program design, implementation, and effectiveness;
- Demonstrate a program's support of the district's mission;
- Justify the costs of a program;
- Determine program strengths and weaknesses;

---

[4] McLeod, S. 2005. "Data-driven teachers." UCEA Center for the Advanced Study of Technology Leadership in Education, p.1. http://dangerouslyirrelevant.org/files/2005_Microsoft_Data_Driven_Teachers.pdf

[5] Marsh, J.A.; Pane, J.F.; and Hamilton, L.S. 2006. "Making Sense of Data-Driven Decision Making in Education." RAND Occasional Paper, p. 6. http://www.rand.org/pubs/occasional_papers/2006/RAND_OP170.pdf

[6] The Center for Comprehensive School Reform and Improvement. 2006. "Program evaluation for the practitioner." Learning Point Associates. http://www.centerforcsri.org/files/TheCenter_NL_June06.pdf

[7] Marynowski, S. 2006. "Best practices guide to program evaluation for aquatic educators." Recreational Boating & Fishing Foundation, p. 2. http://www.rbff.org/uploads/Resources_bestpractices/Best_Practices_Guide_to_Program_Evaluation.pdf

- Measure and explain program performance, outcomes, and impacts;
- Reveal program successes to supporters, funders, and stakeholders;
- Validate or discover effective programming methods; and
- Share information about what works with colleagues and districts.

Effective data collection is paramount to any successful program evaluation. Data collection must be purposeful, meaning that evaluators' efforts should focus on answering questions that are tied to identified needs and goals. A paper on DDDM by the Mid-continent Research for Education and Learning (McREL), a private non-profit corporation located in Denver, Colorado, suggests that "Considering different types of data—for example, demographic, student outcome, perception, and school process data—both alone and in combination over time helps create a more complete view of student achievement."[8]

In addition to thorough data collection, McREL also considers sufficient designated resources and strategies for communicating about the process just as important. A dedicated, cohesive data team can support a district in its improvement endeavors. This team should have the primary responsibility for coordinating data collection, analysis, interpretation, and reporting. Communication on the purpose and results of data analysis to all stakeholders must occur throughout the school year, not just when the school or district's annual report card is released.[9]

Some DDDM strategies for school improvement are better than others. A study of several Milwaukee school districts by the Wisconsin Center for Education Research reveals that using data to support inquiry and inform the instructional mission of schools requires coordinated changes in school processes; data collection and management; the use of analytical tools; and the analytical capacity of school personnel.[10] Table 1 presents a matrix of strategies can be used to assess the adequacy of the DDDM process of a district. McREL attests that schools that take the actions described in the right-hand column are more likely to sustain improvement.

[8] Mid-continent Research for Education and Learning. 2003. "Sustaining school improvement: Data-driven decision making." p. 1. http://www.mcrel.org/pdf/leadershiporganizationdevelopment/5031tg_datafolio.pdf
[9] Ibid., p. 2
[10] Mason, S. 2002. "Turning data into knowledge: Lessons from six Milwaukee public schools." Wisconsin Center for Education Research, University of Wisconsin–Madison, p. 8.
http://www.wceruw.org/publications/workingPapers/Working_Paper_No_2002_3.pdf

Table 1. Assessing the Adequacy of the Data-Driven Decision Making Process

Least Effective	Somewhat Effective	Most Effective
Purposeful Data Collection and Analysis		
Data collection is not aligned with identified needs and goals. Different types of data may be collected, but the focus of data collection is primarily on student outcomes. Data analysis focuses on measures of student achievement over time. Achievement data are disaggregated.	Data collection is aligned with identified needs and goals. Data collection includes several forms of student outcome data but limited amounts of data. Multiple measures of student achievement data are disaggregated and analyzed over time. Different types of data may also be examined but not in combination with other types of data.	The purposes for data collection are clearly stated, and data collection is aligned with identified needs and goals. Appropriate amounts and types of data are collected. Different types of data from a variety of sources, including disaggregated data, are examined over time, alone and in combination.
Resources and Supports		
A data team may be in place, but it is viewed as an ad hoc group rather than a standing committee. Some technology may be available to support the team's work, but no training is provided.	A data team exists and meets on a regular basis. Time is provided for the team to meet. Limited training and technology to support data collection and analysis are available.	Data structures and processes are in place, including a data team, adequate time, appropriate technology, and training. These structures are viewed as permanent, revisited regularly, and revised as necessary.
Communication		
Communications about data are sporadic and intended only as "information dissemination," not for the purpose of discussion and improvement.	Communications about data occur on a regular but limited basis. Communications are mainly for providing information, but there are some opportunities for stakeholders to participate in discussions about data.	There are clear communications about all aspects of data collection, analysis, and use. Communications about data occur on a regular and timely basis. Communications include discussions that provide opportunities for stakeholders to participate in the decision making process.

Source: McREL, 2003

Considerations before Evaluating

With an appreciation for the importance of data to improve school programs, districts may consider other factors about their programs before they begin the process of evaluating them. An initial consideration may be the soundness of a program based on scientific research, as demanded by the No Child Left Behind Act (NCLB) for certain types of programs. Other considerations may include planning-related tasks such as deciding which programs to evaluate and determining how much it will cost to do so.

Scientifically-based research (SBR) can be used to help schools make critical decisions about curriculum and instruction. SBR involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs.[11] NCLB requires educational programs and practices to be based on SBR. While the federal policy impacts practicing educators in the curriculum areas of reading, mathematics, and science, it also impacts instructional strategies, professional development, parent involvement, and all federally-funded programs.[12]

Beyond the narrow scope of SBR as it pertains to NCLB, such evidenced-based research may inform teaching practices, curriculum decisions, and school-wide programs. Guidelines from the U.S. Department of Education stress that educators need to take into account three perspectives when weighing the evidence in favor of adopting a particular program or practice:[13]

- The theoretical base of the reform practice or program;
- The implementation and replicability information; and
- The evidence of effects on student achievement.

Building on these principles, Table 2 on the following page outlines questions that districts may consider in adopting a program or practice to ensure that it is based on scientific research.

[11] Margolin, J., and Buchler, B. 2004. "Critical issue: Using scientifically based research to guide educational decisions." North Central Regional Educational Laboratory. http://www.ncrel.org/sdrs/areas/issues/envrnmnt/go/go900.htm

[12] Ibid.

[13] U.S. Department of Education. 2002. "Scientifically based research and the Comprehensive School Reform Program." p.4. http://www2.ed.gov/programs/compreform/guidance/appendc.pdf

Table 2. Research Consumer Questions for Educational Programs/Practices

General Questions	Judging Quality
Theoretical Base	
What are the ideas behind this practice or program?	Is there a clear, non-technical description of the central idea and goals of the practice or program?
What are its guiding principles?	Is there a clear description of the instructional activities that are central to this program or practice?
How does it work?	Is the practice clearly tied to an established learning theory, e.g. child development or language acquisition?
Why does it work?	
Implementation and Replicability	
Has this program or practice been widely used?	How many schools have used this practice or program?
Where is this reform likely to work?	Did the schools using it fully implement the practice or program?
Under what circumstances is it most effective?	In what settings has it been implemented?
	Has improved student achievement been convincingly demonstrated in a variety of settings?
Evidence of Effects on Student Achievement	
Is there evidence based on rigorous research showing that this practice and/or program improves student achievement?	Are there studies looking at the impact on students of that practice or program?
	Are those studies of high quality
	Are there at least 5 high quality studies?
	Do 4 of the 5 high quality studies show that the practice improves student achievement?
	If yes, are the findings significant in 3 of those 4 studies?

Source: U.S. Department of Education, 2002

Following a district's determination if its current programs are of the highest quality based on scientific evidence, it may then seek to prioritize the evaluations of those programs. Having a complete picture of all the programs is necessary to do this. The New York State Teacher Resource Centers has published a set of instructions for this purpose. The step-by-step process is as follows:[14]

List all of the programs/activities provided.

> Briefly provide names of programs, courses, and resources that your district provides.

Who is a direct participant of the program?

> List people (e.g., science teachers, kindergarten students, parents) who participate in or use the program/resource/activity directly.

Who is affected by the program indirectly?

> List those who may benefit indirectly as a result of participants' use of the program/resource/activity.

With whom will part or all of the evaluation information be shared? (The evaluation 'audience')

> List those who may see or hear about information contained in the evaluation (e.g., community members, state education officials, funders, Board of Education, etc.). Pay particular attention to those who have asked you for evaluation information.

Rate the stakes of this program (e.g., High stakes – a program with high cost, high public visibility, or outcomes are extremely important).

> Consider cost, importance of outcomes and other aspects of the program and rate it as high, medium or low stakes.

For each program/activity/resource decide whether you interested in an evaluation that helps you to improve program components (formative), an evaluation that looks at the success that it has in meeting its goals or both?

> Consider why you are interested in conducting an evaluation.

Is there already documented effectiveness of this program?

> Think of whether or not the program already has documented effectiveness. Perhaps you put a great deal of time and energy into evaluating a certain program last year. Despite the importance of the program, you may want to focus on something else this year. Or perhaps you implemented a program with a proven track record (e.g., evidenced through more than one empirical studies).

[14] Aldrich, S. Op cit., p. 8

The above criteria seek to help a district determine which programs deserve its time and resources. Researchers from the University of Vermont also tackled this dilemma. They advise that those delivering direct services and programs begin down the road of program evaluation by first determining the desired outcomes, activities, and indicators. This step should take place during the planning stages of project development. For evaluation purposes it is essential to identify and document these three items, described as follows:[15]

Outcomes should be consistent with what could reasonably be accomplished and not overly idealistic. They provide a foundation for all subsequent program implementation and evaluation activities, and each of the outcomes will need to be evaluated. Focus outcomes on what can realistically be accomplished within the period of program funding.

Activities are the interventions that a program will provide in order to bring about the intended outcomes. Programs offer all sorts of different activities to address their desired outcomes. For the most part, program activities can be classified as any type of direct service or information that is provided to participants.

Indicators act as the gauge of whether, and to what degree, a program is making progress. A program's progress needs to be examined in two distinct ways:

The quantity and quality of the program activities you are delivering, (commonly referred to as process indicators), and

The quantity and quality of the outcomes that your program is achieving (commonly referred to as outcome indicators).

Before the program evaluation can commence, a barrier that will likely arise for any evaluator during this planning phase is cost. While evaluation does not need to be expensive, it does take time and money to plan an evaluation, collect the right information, and use the results to strengthen a program. One general rule for estimating an evaluation budget is 5 to 10 percent of the total program budget, which includes the value of the time that staff will spend on the evaluation, as well as out-of-pocket costs.[16]

The Evaluation Center of Western Michigan University has developed a budgeting tool for program evaluating. With a mission is to advance the theory, practice, and utilization of evaluation, the Center designed a checklist to assist evaluators think through the many issues that should be considered when developing an evaluation budget. The instrument is divided into several categories of typical cost items for

---

[15] Gajda, R., and Jewiss, J. 2004. "Thinking about how to evaluate your program? These strategies will get you started." Practical Assessment, Research & Evaluation, 9(8). http://PAREonline.net/getvn.asp?v=9&n=8

[16] Hosley, C. 2005. "What will it cost? Who should do it?" Minnesota Office of Justice Programs, p. 1. http://www.ojp.state.mn.us/Grants/Program_Evaluation/Wilder_Tips/3.pdf

conducting evaluations, and questions or statements are used as prompts for users to consider. Table 3 is an adapted version of the checklist.

Table 3. Research Consumer Questions for Educational Programs/Practices

Subcategory	Question(s)
Basic Considerations	
Type of agreement for the evaluation	Will this evaluation be funded by a grant, contract, or cooperative agreement?
Condition of payment	Will payment be on the basis of cost reimbursement or a fixed price?
Funding source	Is the funding source a government agency, private foundation, private sector business/industry, or nonprofit entity?
Funding period	Over what period of time will the budget be applicable? Does this include more than one fiscal year?
Budget contact	Who are the designated contacts for budget matters in the funding organization and in the proposer's organization?
Budget limits	Has a limit for the evaluation budget been established, such as percentage of total budget or maximum amount?
Condition of payment	Will monies from the funding agency be made available as a lump sum, periodically (e.g., monthly, quarterly, or on the basis of submission of deliverables), or reimbursable based on submission of invoices with required documentation?
Pre-award costs	Are costs allowable that are directly related to the evaluation activities but are incurred before the official effective date of the contract?
Personnel Costs	
Personnel types	Have key types of individuals to both administer and conduct the evaluation been identified?
Basis	Will personnel costs be determined on the basis of days of effort, percentage of time, percentage of assigned work load, person hours, or some other factor?
Effort	How much time of each person/type will be required to complete identified tasks?
Unit costs	What is the cost of a time/cost unit for each person?
Consistency of costs	What anticipated increases (merit, cost of living, across the board, etc.) can be anticipated for each evaluation year?
Hiring costs	Will recruitment and hiring costs be required to staff the evaluation?
Support staff	Will support staff services be assigned to the evaluation, or will services (e.g., secretarial support, data entry services, etc.) be "purchased" from a pool?

Student assistance	Have all costs for student workers, as defined by institution/agency, been considered?
Supplies and Materials	
Office supplies	What supplies and materials will be needed to simply conduct the normal operations of an evaluation, e.g., office supplies and institutional products?
Supplies and Materials (cont'd.)	
Specialized materials	What unusual supplies and materials will be needed for this evaluation, i.e., reference books, specialized printing cartridges, evaluation letterhead, data collection articles, etc.? (Note: computer software sometimes is restricted or requires special permission or must be listed separately as a computer cost.)
Communications	
Phone installation	Will any new installation of telephones, computer ports, or other communication equipment be required?
Phone usage	What are the fixed costs for the use of telephones, computers, and other communications equipment that can be identified as chargeable to the evaluation?
Evaluation-specific costs	Other than the usual administrative communications, will there be other costs for large mailings, express services, etc.?
Long-distance costs	What long distance telephone and facsimile charges and other variable communications costs will be incurred?
Postage/mail services	What postage or other forms of mail services will be required, giving special attention to mass mailings of surveys, notices, invitations, etc.?
Copying and Printing	
In-house copying	How many persons will be making in-house copies, and will there be a system to monitor individual usage?
Per-page costs	What is the current and anticipated rate of copying over the long-term evaluation?
Reports	How many copies and of what type/quality must reports be submitted?
Promotional materials	Will the evaluation produce promotional materials, i.e., brochures, pamphlets, etc.?
Data collection	What printing costs will be incurred as a part of the data collection

	process?
Outsourcing	Can all printing jobs be handled inside the organization, or will some outsourcing be required?
Other – Consultants	
Name or type	Who or what types of persons will serve as consultants?
Service	What service will each consultant perform?
Rate and total time	What is the allowable rate or basis for performing the service, e.g., 10 days @ $500/day?
Reimbursable expenses	Will travel and other required expenses incurred by the consultant during provision of service be reimbursable?

Source: Western Michigan University, 2001

While this checklist covers some of the most common evaluation costs, it may be modified or adapted to fit the specific needs of the user. The last category deserves some extra attention, as the use of consultants to gain specialized expertise or to accomplish specific tasks or activities is a common practice for program evaluations. Consultants are often considered to be independent contractors and are not included under personnel costs, and seldom are they provided with fringe benefits or other support services normally available to evaluation employees. Table 4 displays a few considerations that a district may need to take into account when deciding whether or not to hire an external evaluator.

Table 4. Considerations to Hiring an External Evaluator

Questions to Ask	Pros of Hiring	Consequences
Does the current staff have the expertise to conduct an evaluation that meets the district's needs?	Access to specialized knowledge and experience in program evaluation	Increased out-of-pocket costs
Can the staff devote enough time to the evaluation on their own?	Objective reports on evaluation results and implications	Preparation time to select an evaluator and to acquaint the evaluator with your program
Is receiving external, objective assistance and feedback important to stakeholders?	Increased credibility of the evaluation results	Possible skepticism or resistance among program staff if the evaluator is seen as an outsider

Source: Minnesota Office of Justice Programs, 2005

Hiring an external evaluator from a research institute or a consulting firm is one option for a district. Other options include using an in-house evaluation team supported by an outside consultant and/or program staff. For this, an in-house evaluator would serve as the team leader—guiding the development of the evaluation design, conducting data analyses, and selecting or developing questionnaires—and be supported by both program staff and the outside consultant. Regardless of the scenario pursued, a decision to hire evaluators or to contract for their services should be governed by a desire to maximize several values:[17]

- The technical skills of the evaluators;
- The evaluator's familiarity with the details of education programs;
- The disinterestedness/independence of the evaluator; and
- The utility of the evaluation for the decision makers.

With a solid understanding of the district's programs and measurable goals as well as the potential costs to evaluate them, district leaders are then challenged with designing and implementing a program evaluation. Whether the evaluation is conducted in-house or in conjunction with an outside consultant, applying a consistent methodology leads to clearer judgments about how well the program's objectives have been achieved. A framework for program evaluating is discussed below.

Program Evaluation Framework

Evaluation involves the systematic and objective collection, analysis, and reporting of information or data. Using the data for improvement and increased effectiveness then involves interpretation and judgment based on prior experience. The purpose of evaluating a program is to essentially determine its worth. To answer questions about program impacts and processes, an evaluation should be conducted as a systematic assessment. Evaluations should follow a systematic and mutually agreed on plan. Plans will typically include the following:[18]

Determining the goal of the evaluation

What is the evaluation question?

What is the evaluation to find out?

How the evaluation will answer the question

What methods will be used?

Making the results useful

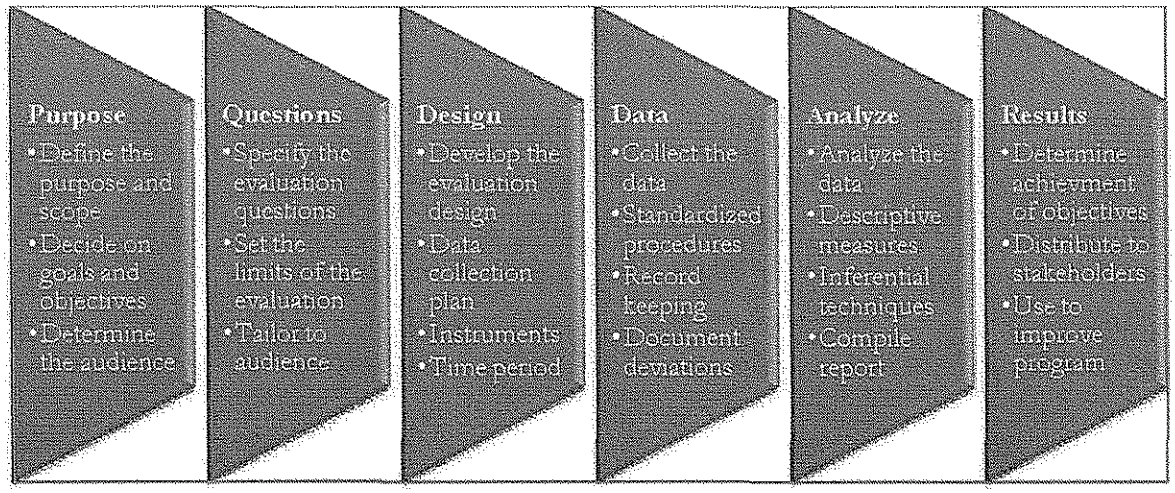
[17] National Institute of Justice. 1992. "Evaluating drug control and system improvement projects: Guidelines for projects supported by the Bureau of Justice Assistance." p. 13.
http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/chapter_4_nij_guide.htm

[18] Shackman, G. Op. cit.

How will the results be reported so that they can be used by the organization to make improvements?

The evaluation process can be described as involving six progressive steps. These steps are shown in Chart 1, each of which are discussed in greater detail beginning on the next page.

Chart 1. Overview of the Evaluation Process



Source: Development Associates, Inc., 1996

The first step of planning is to define an evaluation's purpose and scope, which helps set the limits of the evaluation. Then the audience for the evaluation must be identified, which may include school administrators, planners, and local (or state) decision makers. While the goals of the evaluation are more easily determined based on whether the program is new or established, defining the scope depends on the evaluation's purpose and the information needs of its intended audience. More specifically:[19]

> These needs [of the audience] determine the specific components of a program which should be evaluated and on the specific project objectives which are to be addressed. If a broad evaluation of a curriculum has recently been conducted, a limited evaluation may be designed to target certain parts which have been changed, revised, or modified. Similarly, the evaluation may be designed to focus on certain objectives which were shown to be only partially achieved in the past. Costs and resources available to conduct the evaluation must also be considered in this decision.

Following a determination of the evaluation purpose, a good way to begin formulating evaluation questions is to carefully examine the program objectives; another source of questions is to anticipate problem areas of the program.[20] Questions establish boundaries for the evaluation by stating what aspects of the program will be addressed. Negotiating and prioritizing questions among stakeholders further refines a viable focus for the evaluation. The development of evaluation questions consists of several steps:[21]

- Clarifying the goals and objectives of the project;
- Identifying key stakeholders and audiences;
- Listing and prioritizing evaluation questions of interest to various stakeholders; and
- Determining which questions can be addressed given the resources and constraints for the evaluation (money, deadlines, etc.)

Evaluation methods should be selected to provide the appropriate information to address stakeholders' questions. When designing an evaluation, program evaluators can use a variety of types of evaluation, such as for needs assessments, formative,

---

[19] Fleischman, H.L., and Williams, L. 1996. "An Introduction to program evaluation for classroom teachers." Development Associates, Inc. http://teacherpathfinder.org/School/Assess/assess.html

[20] Ibid.

[21] Mid-Continent Comprehensive Center. n.d. "Developing evaluation questions." http://www.mc3edsupport.org/community/knowledgebases/developing-evaluation-questions-820.html. From National Science Foundation. 1997. *User-Friendly Handbook for Mixed Method Evaluations.*

summative, process, outcomes, and the like.[22] The type of evaluation undertaken to improve a program depends on what the evaluator wants to learn about the program. The overall goal in selecting an evaluation method is to get the most useful information to key decision makers in the most cost-effective and realistic fashion. Consider the following questions:[23]

- What information is needed to make current decisions about a program?
- Of this information, how much can be collected and analyzed in a low-cost and practical manner, e.g., using questionnaires, surveys and checklists?
- How accurate will the information be?
- Will the methods get all of the needed information?
- What additional methods should and could be used if additional information is needed?
- Will the information appear as credible to decision makers, e.g., to funders or administrators?
- Will the nature of the audience conform to the methods, e.g., will they fill out questionnaires carefully, engage in interviews or focus groups, let you examine their documentations, etc.?
- Who can administer the methods now or is training required?
- How can the information be analyzed?

If the answers to the evaluation questions are to be reliable and believable to program stakeholders, the evaluation must collect information in a consistent and thoughtful way.[24] The data should be recorded carefully so they can be tabulated and summarized during the analysis stage, and deviations from the data collection plan should be documented so that they can be considered in analyzing and interpreting the data.[25] The collection of information can involve individual interviews, written surveys, focus groups, observation, or numerical information such as the number of participants. Table 5 provides an overview of the major methods used for collecting data for evaluations.

---

[22] Fleischman, H.L., and Williams, L. Op. cit.
[23] McNamara, C. 2002. "Basic guide to program evaluation." Authenticity Consulting, LLC. http://managementhelp.org/evaluatn/fnl_eval.htm
[24] W.K. Kellogg Foundation. 1998. "Evaluation Handbook." p. 14. http://www.ojp.usdoj.gov/BJA/evaluation/links/WK-Kellogg-Foundation.pdf
[25] Op Fleischman, H.L., and Williams, L. Op. cit.

Table 5. Overview of Methods to Collect Information

Method	Overall Purpose	Advantages	Challenges
Questionnaires, Surveys, Checklists	To quickly and/or easily get lots of information from people in a non threatening way	•Can complete anonymously •Inexpensive to administer •Easy to compare and analyze •Administer to many people •Can get lots of data •Many sample questionnaires already exist	•Might not get careful feedback •Wording can bias target's responses •Are impersonal •In surveys, may need sampling expert • Does not get full story
Interviews	To fully understand someone's impressions or experiences, or learn more about their answers to questionnaires	•Get full range and depth of information •Develops relationship with target •Can be flexible with target	•Can take much time •Can be hard to analyze and compare •Can be costly •Interviewer can bias target's responses
Documentation Review	To gather impressions of how the program operates without interrupting the program; is from review of applications, finances, memos, minutes, etc.	•Get comprehensive and historical information •Does not interrupt program or target's routine in program •Information already exists •Few biases about information	•Often takes much time •Info may be incomplete •Need to be quite clear about what looking for •Not flexible means to get data; data restricted to what already exists
Observation	To gather accurate information about how a program actually operates, particularly about processes	•View operations of a program as they are actually occurring •Can adapt to events as they occur	•Can be difficult to interpret seen behaviors •Can be complex to categorize observations

Method	Overall Purpose	Advantages	Challenges
			•Can influence behaviors of program participants •Can be expensive
Focus Groups	To explore a topic in depth through group discussion, e.g., about reactions to an experience or suggestion, understanding common complaints, etc.;	•Quickly and reliably get common impressions •Can be efficient way to get much range and depth of information in short time • Can convey key information about programs	•Can be hard to analyze responses •Need good facilitator for safety and closure •Difficult to schedule 6-8 people together
Case Studies	To fully understand or depict target's experiences in a program, and conduct comprehensive examination through cross comparison of cases	•Fully depicts target's experience in program input, process and results •Powerful means to portray program to outsiders	•Usually quite time consuming to collect, organize and describe •Represents depth of information, rather than breadth

Source: Authenticity Consulting, LLC, 2002

Analyzing the collected data involves tabulating, summarizing, and interpreting the data in such a way as to answer the evaluation questions. The timing of the data analysis and interpretation should be driven by the evaluation questions. For example, evaluators might decide to interpret some data formatively so they can see what implementation modifications and adjustments are suggested.[26] Formative evaluation is designed to help the program confirm its directions, influence, or help to change it; summative evaluation summarizes the whole process and describes its destination.[27] Table 6, below, describes formative and summative evaluations more closely.

[26] Marynowski, S. Op cit., pp. 6-7
[27] Pratt, N. 2004. "Evaluation research in education." Univeristy of Plymouth (UK). http://www.edu.plymouth.ac.uk/resined/evaluation/index.htm

Table 6. Formative vs. Summative Evaluations

Formative Evaluation	Summative Evaluation
Examines the program design, technology, delivery, content, personnel, procedures, and inputs	"Summarizes" a program by describing what happens after delivery of the program and looks at whether a program is meeting its stated objectives
Helps to define the scope of a program and to identify appropriate goals and objectives	Provides information about whether a program reached the intended target audience and whether the participants found the program helpful or useful
Can be used to test ideas and strategies before a program is designed	Seeks to determine whether the program itself caused the observed outcomes, whether there were secondary or unexpected program outcomes, and the relative costs and benefits of the program
Can be done while the program is in progress, to determine if the program is on the right track, providing information for fixing weaknesses, correcting shortcomings, or dealing with unforeseen obstacles in program delivery	Provides ideas for future modifications or improvements in your programs
Provides definite information to create a well-designed and well-targeted program from the start.	Supplies unbiased information for discussing (or defending,) the impacts, benefits, and cost-effectiveness of your program with administrators, funders, sponsors, community members, and other stakeholders

Source: Recreational Boating & Fishing Foundation, 2006

Data tend to fall into two categories of information: quantity and quality. Typically, quantity measures—used to evaluate both process and outcome indicators—are numerical descriptions of program activities and achievements, while quality measures—determine and document the effectiveness of the program's activities and services—often portray program activities and achievements through narrative descriptions.[28] Examples of tools for quantitative data include surveys, performance assessments, and content analyses. Those for qualitative data typically include observations, interviews, and focus groups. Table 7, below, describes the appropriate use of quantitative and qualitative data.

[28] Gajda, R., and Jewiss, J. Op cit.

Table 7. Quantitative vs. Qualitative Data

Quantitative	Qualitative
Evaluate large-scale programs	Collect descriptive information to understand attitudes, beliefs, and perceptions
Generalize results to large populations	Recognize program outcomes or impacts
Measure levels of knowledge, attitudes, beliefs, perceptions, or change in behavior	Measure and understand behavioral change
Measure the amplitude of program outcomes or impacts, or causes and effects	Judge the nature of causes and effects
Determine if changes are statistically significant	Understand complex issues and program context
Compare or rank features of various groups	Identify unintended or unexpected program outcomes

Source: Recreational Boating & Fishing Foundation, 2006

After analyzing the collected data, evaluators can make a judgment about how well the goals of the program have been met and then reporting their findings to stakeholders who have an interest in the program and its impacts. The level and scope of content depends on to whom the report is intended. For example, administrators need general information for policy decision making, while teachers may need more detailed information which focuses on program activities and effects on participants. The report should cover the following:[29]

- The goals of the evaluation;
- The procedures or methods used;
- The findings; and
- The implication of the findings, including recommendations for changes or improvements in the program.

The framework for conducting a program evaluation gives a sense of what is involved in the process, from asking the right questions to evaluation tools and analyzing collected data. Following these basic steps can provide a solid base from which to make decisions that ultimately lead to stronger programs. The evaluation overview provided in this section is complimented by examples of districts' methods of

[29] Fleischman, H.L., and Williams, L. Op cit.

evaluation in Section Two. Hanover gathered the examples from publicly available sources and interviewed school leaders for information about their district's evaluationprocess.

Section Two: Survey of Districts' Practices in Evaluating Programs

Section One of this report presented guidelines to planning a program evaluation and a general framework for program evaluating in the context of education. This section identifies program evaluation practices of public school districts through discussions with school leaders. It begins with a discussion of the methodology to select the districts that Hanover reached out to for interviews. To supplement the interviews, we also observed examples of evaluation protocols that are publicly available from districts' websites.

Methodology for Selecting Districts

Hanover contacted 17 public school districts in 12 states to request an interview with the director of each district's department of evaluation or other similar office. The districts were primarily chosen based on type, size, and locale as indicated by data from the National Center of Education Statistics (NCES).[30] With one exception, the following criteria were selected to determine the contacted districts:

- **Type** – Regular School District
- **Locale** – City/Suburban: Large
- **Total number of schools** – Greater than or equal to 50
- **Total number of students** – Greater than 25,000

The selected schools also were chosen for having prominent research/evaluation departments. A search of school districts' websites for those that met the above criteria led us to select the following districts:

- Denver Public Schools (CO)
- Broward County Public Schools (FL)
- Gwinnett County Public Schools (GA)
- Chicago Public Schools (IL)
- Boston Public Schools (MA)
- Prince George County Public Schools (MD)
- Baltimore County Public Schools (MD)
- Anne Arundel County Public Schools (MD)

[30] "Public School District Search." NCES, Institute of Education Sciences (U.S. Department of Education). http://www.nces.ed.gov/ccd/districtsearch/

- Detroit Public Schools (MI)
- Omaha Public Schools (NE)
- Portland Public Schools (OR)
- Providence Public School District (RI)
- Austin Independent School District (TX)
- San Antonio Independent School District (TX)
- Norfolk Public Schools (VA)
- Loudoun County Public Schools (VA)

The one exception to our methodology was our decision to include Arlington County Public Schools (VA). It is a 'Regular School District' like the others also with a significant evaluation department, but is considered a 'Mid-size' city district comprised of only 33 schools with less than 19,000 students enrolled according to NCES data. We contacted Arlington after viewing its Office of Planning and Evaluation website and numerous online evaluation documents including a framework for accountability and evaluation.

Hanover spoke with seven directors of evaluation departments from the above list of districts, including Arlington County, Baltimore County, Broward County, Gwinnett County, Loudoun County, Portland City, and San Antonio District. The other districts have either not responded to our request, declined to be interviewed, or expressed interest but have not scheduled an interview with us. The remainder of this report is divided into subsections of surveyed districts with which we had contact.

Broward County Public Schools (Florida)

The county is a large suburban region of Fort Lauderdale that serves approximately 259,000 students. The school system has an established Office of Research, Development & Assessment, of which the Research Services Department is responsible for the evaluation of school programs and more. According to the Department website, the activities of the research division include the following:[31]

- Procurement of appropriate data to aid the Superintendent and Administration in decision making;
- Provision of information on contemporary educational issues;

[31] Broward County Public Schools (BrCPS). "Department of Research Services mission." http://www.broward.k12.fl.us/research_evaluation/Mission.htm

- Development, selection, and utilization of appropriate assessment tools, including surveys, for the evaluation of District programs;
- Annual administration of customer surveys to students, parents, and teachers, including the production of school-by-school and District-wide reports;
- Design and execution of studies to evaluate processes and outcomes associated with educational programs and instructional strategies;
- Monitoring the collection and organization of longitudinal data to determine trends in student achievement and changes in demographics over time;
- Provision of assistance to schools and departments for interpretation of data and evaluation results;
- Reporting mandated data requests from the Florida Department of Education, U.S. Department of Education, and other requests from outside organizations and individuals; and
- Evaluating proposed research collaborations with outside agencies and individuals engaging in research activities within the District.

Hanover interviewed the Associate Superintendent responsible for the Office to find out more information. We were informed that the Office is funded through a general allocation, the amount of which is determined by an annual budget review process. The Research Services Department uses a variety of report formats to publish evaluation findings including formal evaluation reports, program status reports, research briefs, information briefs, and data analyses. Evaluations are generally conducted by independent consultants and are more comprehensive in scope.

Much of the evaluations completed encompass Title I programs that require annual evaluations and other grant-funded programs. Otherwise, program evaluation requests typically come from the school board, which target new programs, major expansions of current programs, and high-cost programs. Generally, program evaluations may be formative or summative in nature or a combination of both. Most of the evaluating is conducted in-house with a large staff of research specialists, evaluation administrators, database researchers, and technicians. All staff have advanced training in research, evaluation, and measurement methodology.

The Office of Research, Development & Assessment operates on a five-year calendar, the most recent cycle of which began in April 2009.[32] It has established "Guidelines and Procedures for the Annual Evaluation Process" as part of the

[32] BrCPS. 2009. "Research Development & Assessment five-year research & program evaluation calendar." http://www.broward.k12.fl.us/research_evaluation/consultantpage/RFP/5-YearCalendar4-16-09.pdf

Research Services Department's process to survey executive leaders for projects or programs that require annual evaluations, reports, or analyses to be completed during the subsequent five school years.[33] Leaders must complete and submit an electronic request form for each project they want to add to the five-year calendar or any project currently on the calendar that they want to modify or delete. With regard to the funding of evaluations:[34]

> Evaluation costs are determined by the scope of the project. When submitting a project for inclusion on the Five-Year Calendar, the amount and source of the funds for all evaluation activities must be included on the Request Form. … However, projects requiring District support have no guarantee that funds will be available. District funds and staff are limited. If necessary, Executive Leaders will prioritize projects requested for District funding each year. The extent to which these projects are funded will be determined by available funds and staff capacity of the Research Services Department.

Concerning specific projects, the Office annually conducts a "customer satisfaction survey" for community feedback; Department of Children and Families (DCF) Interagency Agreement annual report on educational services to children in the foster care system of Broward County; research brief on promotion/retention district initiative; and various Title I programs. Other evaluations include Early Reading First, Smaller Learning Communities, and Teaching American History.

Gwinnett County Public Schools (Georgia)

A large suburb of metropolitan Atlanta, Gwinnett County serves over 156,000 students. The school system's Department of Research & Evaluation provides direct support in accelerating school and system-wide improvement through the following activities:[35]

- Provides technical support for the Research-Based Evaluation System (RBES);
- Support for the development of data systems;
- Management of the instructional research approval process; and
- Research consultation to schools, departments, and professional staff.

An interview by Hanover with the Executive Director of the Department revealed that it operates on a fixed budget of $618,000 ($143,000 excluding staff), of which 65

[33] BrCPS. 2008. "Guidelines and procedures for the annual evaluation process." p. 1. http://www.broward.k12.fl.us/research_evaluation/consultantpage/RFP/5yrCalGuidelinesProceduresv1-24-08.pdf
[34] BrCPS (2008). Loc. cit.
[35] Gwinnett County Public Schools (GCPS). "Department of Research and Evaluation." http://www.gwinnett.k12.ga.us/gcps-mainweb01.nsf/pages/OfficeofResearchandAccountability

percent is specifically for evaluating purposes. The Department's primary focus is running the RBES, an accountability system for improving schools systematically measures a school's progress. With this system, every school in the county receives an annual report card.

Beyond this, the Department operates on an informal basis and has a limited focus due to small resources. It administers perception surveys to gather feedback on school programs from selected students, parents, and staff. Requests for data from teachers and principals also are common. Program directors also rely on the Department for data for their own projects. The school board encourages the conduct of well-designed educational research projects within the district, and outside evaluators are rarely used.

However, the district does appear to have an interest in participating in external research studies. The district has a list of suggested research questions/topics in the areas of accountability and assessment; English language learners; facilities and operations; foreign language; health services and school social workers; mathematics; science; and special education and psychological services.[36] According to the Department's website, "the Executive Director of Research and Evaluation must provide written approval before a research project may be conducted in the district. The Director and the appropriate division will be responsible for monitoring any approved research."[37]

Baltimore County Public Schools (Maryland)

This district is located in a large suburban area serving over 104,000 students. The Department of Research, Accountability & Assessment is responsible for the development, implementation, and evaluation of system research activities, including all aspects of conducting and reporting results of research and program evaluations related to the many factors that impact student achievement, and analyzing and reporting performance results.[38]

The Office of Research designs and implements evaluations intended to provide information about the efficacy of selected programs. It also provides research and statistical analysis services to the schools and offices of the district as well as support in the design and analysis of survey research. In addition, the Office is responsible for reporting official statistical information for the district.[39]

---

[36] GCPS. 2009. "Research questions/topics of interest." http://www.gwinnett.k12.ga.us/gcps-mainweb01.nsf/05A0BAC46B67ACE88525762700689BB8/$file/2009-10_Suggested_Studies_2009_09_04.pdf
[37] GCPS (n.d.). Op. cit.
[38] Baltimore County Public Schools (BaCPS). "Research, Accountability, and Assessment: Mission statement." http://www.bcps.org/offices/accountability_research_testing/
[39] BaCPS. "Research, Accountability, and Assessment: Office of Research." http://www.bcps.org/offices/accountability_research_testing/research.html

Through an interview with the Executive Director of the Department, Hanover learned that all instruction and instructional support programs undergo a cycle of evaluation every five years. Other determinants of program evaluation include the information needs of the superintendent based on strategic goals and requests from departments, which are brought to the attention of the superintendent. The school board does not play a major role in prioritizing evaluations but instead primarily sets policy for the evaluation of programs. Under this policy, the evaluation process will:[40]

- Assess the applicability of methods, procedures, materials, and theories as appropriate and specific to differences in populations and circumstances;
- Assess the effectiveness of a program in accomplishing its original goals and objectives;
- Assess the appropriateness of measurable outcomes and their link to achievement at all levels;
- Report on the success of the program in increasing achievement at appropriate levels and opportunities for improvement of the program; and
- Suggest changes in goals and objectives as appropriate.

According to the Executive Director, the Department is well-funded with a $2 million budget (excluding salaries) and has a sizable enough staff of professionals—many with doctorate degrees in research methodology—to be able to do much of the evaluating in-house. He also credits the district's dedication towards (i.e., funding for) accurate data storage for the Department's ability to effectively evaluate programs. The exception to internal evaluating is mainly programs funded by grant money that require the use of external evaluators.

The Department is responsible for designing and carrying out evaluations. Research designs use mixed methodologies and are quasi-experimental (i.e., empirical approach lacking random assignment). Historical data captured in the district's data warehouse enables longitudinal studies. The timeframes for evaluations, commonly in months, are determined by the superintendent's needs. Specific budget line items for evaluations are not the norm; rather, the Department receives approval for expenditures from the superintendent.

Portland Public Schools (Oregon)

[40] BaCPS. 2009 "Evaluation of the instructional program." School Board Policy 6501, p.1. http://www.bcps.org/system/policies_rules/policies/6000Series/POL6501.pdf

The school system is classified as a large city locale with about 46,000 students. The district's Department of Research, Evaluation, & Assessment has six main functions, including:[41]

- Administer district-wide testing programs and other tests used by a large number of schools as well as provide reports of results to a wide variety of audiences;
- Monitor student completion of work samples, early childhood literacy assessments, common literacy assignments and other local assessments;
- Produce reports for school administrators summarizing course grades;
- Conduct program evaluations, including reports required by external funding sources as well as internally identified programs about which we want objective analyses of implementation and/or impact;
- Provide support to schools and departments that want to conduct paper-based or electronic surveys; and
- Implement School Board policy by reviewing all requests to conduct research with students or staff in Portland Public Schools.

We spoke with the Director of the Department to learn more. The Director noted the Department's small staff and its very informal operations. The Department operates on a fixed budget of about $1.5 million, a large portion of which goes to staffing costs. It conducts primarily internal assessments for accountability purposes, including test results, enrollment reports, and school profiles.

While the Department receives mostly one-time requests for data from individual schools such as test scores, it does conduct larger evaluations from time to time, for example the teaching of high school algebra to eighth graders. The evaluation of grant-funded programs is different, such that the district uses external evaluators for them. Contracted for about two-thirds of all grant program evaluations, the firms are carefully chosen based on qualifications of the targeted area of assessment.

Regarding in-house evaluations, the district is committed to the importance of collecting and using program evaluation data at all levels of the organization. Typical evaluation tools include interviews, focus groups, and observations. The most recent publicly available evaluation reports are posted on the Department's website, which include topics of high school reform and 21st Century Community Learning Centers.[42]

---

[41] Portland Public Schools (PPS). "Department of Research, Evaluation, and Assessment." http://www.pps.k12.or.us/departments/research-evaluation/index.htm

[42] PPS. "Evaluation Reports." http://www.pps.k12.or.us/departments/research-evaluation/1512.htm

San Antonio Independent School District (Texas)

San Antonio ISD is a large city school system that serves nearly 55,000 students. Its Accountability, Research, Evaluation & Testing Department helps guide academic instruction, data management, and administrative decision making. Specifically, the School Planning & Evaluation Office is committed to conducting in-depth formative and summative evaluations of District initiatives.[43]
In an interview with the Senior Director for School Planning and Evaluation Office, Hanover learned that it is a small collaborative division comprising primarily the Senior Director, a Director for Testing, Coordinator for Institution & Community-Based Research, Coordinator for Accountability & Compliance. Community-based research involves outside surveys for feedback on school programs.

The Senior Director noted that full program evaluations are less common unless required, such as state compensatory education programs or charter school reports. Rather, the Office primarily receives requests for data from teachers and departments that include brief analyses and charts. A data warehouse has been built over the years to facilitate longitudinal studies by the Office.

When full evaluations are necessitated, the Senior Director commented that the Office follows the CIPP (Context, Input, Process, Product) evaluation model. In general, these four parts of an evaluation respectively ask, "What needs to be done? How should it be done? Is it being done? Did it succeed?"[44] She cited the current District Awards for Teacher Excellence (DATE) program that allows districts to create or continue a system of awards for educators who demonstrate success in improving student achievement.

Larger reports published by the Office are typically for school board requests or those of district administrators. The district's Annual Performance Review (APR) was cited as an example by the Senior Director. The APR explains the ratings on the district by the state and covers academic performance indicators by schools and the district as a whole. Quarterly reports on grades and attendance also are standard.

External evaluators are contracted in some cases, according to the Senior Director. These evaluators are typically used when only when required, such as for the assessment of a grant-funded program—DATE falls in this category. However, the Office still plays a role in getting the external evaluators the data that they need. The same applies to departments who occasionally contract out evaluations.

Arlington Public Schools (Virginia)

---

[43] San Antonio ISD. "Accountability, Research, Evaluation & Testing: Our mission." http://www.saisd.net/dept/aare/
[44] Stufflebeam, D.L. 2002. "CIPP Evaluation Model Checklist." The Evaluation Center, Western Michigan University. http://www.wmich.edu/evalctr/checklists/cippchecklist.htm

Arlington is a mid-sized city school district that serves about 18,000 students. The Planning & Evaluation Office is responsible for administering, supervising and/or coordinating the testing program, research, strategic plan, division and school management plans, accreditation, attendance and numerous state and local reports. A major function of the Office is to produce annual academic performance reports and use data from scoring contractors and analyses conducted by staff. Regarding the evaluation duties of the Office, it is responsible for the following:[45]

- Prepares the annual report on progress on the strategic plan indicators to the School Board;
- Facilitates the process for summative evaluations of instructional programs and departments for the purpose of continuous improvement;
- Ensures that summative evaluations follow standard practices as outlined by the district's framework for evaluating;
- Coordinates data collection and reporting for summative evaluations of instructional programs and departments;
- Validates findings presented in summative evaluation reports prepared by instructional programs and departments;
- Provides assistance and support to senior staff in the development of annual department plans;
- Conducts special studies and evaluations requested by the Superintendent;
- Coordinates and reports on the biennial Community Satisfaction Survey;
- Approves surveys administered within the district;
- Provides expertise on survey design, questionnaire development, administration and reporting for surveys conducted within the district;
- Manages the 360° performance evaluation process for administrators, and provides the guidelines and support for the development of administrator work plans; and
- Supports the work of and incorporates feedback from the Advisory Committee on Accountability and Evaluation.

As for the Committee referred to in the last point, it reviews progress on accomplishment of the accountability and evaluation system and advises the

[45] Arlington Public Schools (APS). "Department of Information Services Office of Planning & Evaluation." http://www.apsva.us/154010829233400/site/default.asp

superintendent on programs, practices, and behaviors related to its implementation.[46] Specific tasks of the Committee—comprising parents, other community members, students, teachers, and other staff—are to review evaluation requirements and review progress and data on implementation of the system to evaluate progress on instructional programs and other school plans.

The district follows a framework for the systematic evaluation of personnel, schools, and programs, entitled "A Framework for Systematic Accountability and Evaluations."[47] This document describes the district's beliefs and expectations concerning evaluation, such that it will evaluate and report progress through the results of work on 'Annual Priorities' set by the school board. The school system must report qualitative and quantitative evaluation criteria to assess progress and analyze results, both short term and over time.

The Framework also outlines the expected methods for and components of evaluation. Two models of evaluation are required to provide the information used to assess progress: (1) annual formative evaluations to produce information useful for improving program and service implementation and for revising current school, program and staff practices; and (2) periodic summative evaluations to address the degree of goal achievement, and the need for continuation, revision or termination of programs and services. Specific to summative evaluation, the following are the activities in the six years of the instructional program:[48]

Year 1 – Implementation and staff development on new text and materials; Evaluation design and preliminary data collection
Year 2 – Refinement of evaluation design and data collection
Year 3 – Evaluation Year/Data interpretation, report and recommendations
Year 4 – Program revisions and implementation
Year 5 – Implementation and materials need identification
Year 6 – Textbook and materials adoption

The Assistant Director of Planning and Evaluation confirmed to Hanover in an interview that the district has a six-year evaluation cycle but mentioned that is somewhat flexible and adjustments are made occasionally. The district developed an evaluation schedule that outlines the curriculum or program areas and departments that it will evaluate in a given year in conjunction with textbook adoption years. For example, the Health & Physical Education (with Driver Education) and Career, Technical & Adult Education programs were evaluated during the 2008-09 school

[46] APS. 2007. "Advisory Committee on Accountability and Evaluation." http://www2.apsva.us/15401086155450550/site/default.asp?
[47] APS. 2007. "A Framework for systematic accountability and evaluation." p. 1. http://www.apsva.us/154010829233400/lib/154010829233400/Framework%20REV%20Aug%2007.pdf
[48] Ibid., pp. 7-8

year. Both of these programs will be reevaluated in 2014-15 according to the online calendar.[49]

Before commencing a program evaluation, a research design is typically submitted to the school board for feedback. Instructional committees also have advisors who report to the board. This takes place the year prior to the data collection year to ensure that the evaluation is asking the right questions and so that adjustments can be made as needed before implementation.

To carry out a program evaluation, the Office works in collaboration with the Department of Instruction on a research design which is guided by the evaluation Framework (discussed above) and in the board's Policy and Procedures for Accountability and Evaluation. According to the Assistant Director, any research design seeks to answer questions such as the degree to which the district implemented the program as intended, the outcomes for intended recipients, and overall levels of satisfaction with the program.

Concerning the methodology of evaluations, classroom observation is a common component used by the Office. The district also uses standardized/local assessment measures for outcomes evaluations. The Assistant Director noted that existing tools are used whenever possible rather than creating new ones. Written largely by the Office, final evaluation reports are submitted to the board and typically posted online. The Assistant Director acknowledges that the results impact the professional development of teachers.

The internal staff engaged in evaluating consists of the Assistant Director, a data evaluation specialist, and a support staff member. While this team conducts in-house evaluations, external evaluators are occasionally used (often in combination with internal resources), such as to facilitate focus groups of parents in pre-K evaluation. On average, less than half of the staff's time is spent on program evaluation tasks. Money set aside for program evaluating and other activities of the Office currently totals about $190,000 (annual budget line item), which is used primarily to fund external evaluators.

Loudoun County Public Schools (Virginia)

Classified as a large suburban district, Loudoun County serves nearly 54,000 students. The district's Research Office supports administrators, staff, and teachers with data and research to better inform decisions at all levels throughout the school division. According to its website, the Research Office is responsible for the following:[50]

---

[49] APS. 2007. "Evaluation Schedule."
http://www.apsva.us/154010829233400/blank/browse.asp?A=383&BMDRN=2000&BCOB=0&C=54551
[50] Loudoun County Public Schools (LCPS). "Services of the Research Office."
http://cmsweb1.loudoun.k12.va.us/50910052783559/site/default.asp

- Collecting, analyzing, and reporting data regarding student achievement as requested to the School Board, central office staff and school administrators;
- Managing all aspects of the web-based reporting tool used by administrators and teachers to disaggregate data regarding assessment, attendance, enrollment, and performance;
- Providing assistance with the design and implementation of programs administered within the school division;
- Planning and implementing program evaluations in collaboration with central office staff;
- Assisting school administrators with the identification and measurement of efforts planned for school improvement, including division-wide surveys;
- Designing and analyzing surveys used by school administrators and central office staff to assess effectiveness and improve programs; and
- Producing a quarterly newsletter, issue briefings, and literature reviews that synthesize current research on topics of interest to district administrative and instructional staff.

Hanover had the opportunity to interview the district's Director of Research. He works with a staff of five that use collaborative processes to conduct program evaluations, research reviews, and data collection, and analyses. A specific budget for the Office was not given, but the Director mentioned that about 10 to 15 percent of the budget is dedicated to evaluating or similar tasks. He spends about 20 percent of his time on evaluation-related activities.

Concerning the prioritization of program evaluations, the Director informed us that any programs tied to grants receive top priority. Besides those state- or federally-funded programs, new initiatives and curriculum are evaluated typically over the first three years. The school board occasionally seeks information on particular programs, but requests primarily come from program managers, which are then submitted to assistant superintendents. Curriculum evaluations follow a textbook adoption cycle.

The Office uses a collaborative model to implement program evaluations. Research staff work with program managers to craft evaluation strategies for their programs. A foundational study is put together by research staff based on input from program managers, such as research questions, conditions, indicators, etc. Findings are presented to senior staff to determine the feasibility of proceeding with an evaluation. If approved, the time period for an evaluation is determined by program managers and the resources available to them.

Evaluations are mostly conducted internally except for grant-funded programs that require the use of external evaluators. Common research designs include formative evaluations, case studies, and quasi-experimental methods. The Director referred to the designs as "exploratory," being ad hoc rather than systemic. Data collection tools for process evaluation tend to be surveys, observations, and focus groups, while standardized measures and student achievement data such as the Virginia Standards of Learning are analyzed for judging program outcomes.

Project Evaluation Form

The Hanover Research Council is committed to providing a work product that meets or exceeds member expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.

http://www.hanoverresearch.com/evaluation/index.php

Note

This brief was written to fulfill the specific request of an individual member of The Hanover Research Council. As such, it may not satisfy the needs of all members. We encourage any and all members who have additional questions about this topic – or any other – to contact us.

Caveat

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties which extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of The Hanover Research Council or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every member. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, The Hanover Research Council is not engaged in rendering legal, accounting, or other professional services. Members requiring such services are advised to consult an appropriate professional.

Addendum – District Program Evaluation Practices

Prepared for Madison Metropolitan School District

In this briefing, The Hanover Research Council provides a summary of written responses by the Austin Independent School District to a questionnaire regarding the district's program evaluation process. This briefing is intended to serve as a supplement to our updated May 2010 report. 2 DISTRICT ADMINISTRATION PRACTICE

Addendum

Austin Independent School District (Texas)

Austin Independent School District (AISD) is a large city district that serves about 83,000 students in the State of Texas. The district's Department of Program Evaluation (DPE) within the Office of Accountability works with program staff throughout the district to design and carry out formative and summative program evaluations, particularly of federal, state, and locally-funded programs. According to the AISD website, DPE is responsible for the following:[1]

- Developing an annual agenda that identifies programs to be evaluated and services to be provided;
- Reporting objectively on program implementation and outcomes to program staff, decision makers, and planners in the district;
- Coordinating research requests with external agencies, such as universities and governmental organizations;
- Conducting an annual survey of its employees covering key issues of importance to district administrators; and
- Handling internal data requests and ad hoc research projects for district administrators and decision makers.

Regarding the first point above, the DPE develops a 'Plan of Work' to describe the scope of the department's tasks for the coming year. The annual document identifies programs to be evaluated and serves as the blueprints for evaluation staff to follow throughout the year. As noted in the most recent version, "Evaluation plans are developed through an interactive process involving evaluation and program staff, the executive director of Accountability, and for the coming year, other executive-level district staff."[2]

The DPE's 2009-10 Plan of Work demonstrates the department's structure capabilities. The organizational chart displayed in the document shows that the DPE is led by a Director of Program Evaluation who oversees five teams of evaluation analysts, each headed by an administrative supervisor. Nearly all of DPE's staff hold doctorate degrees. Over two dozen evaluation projects were planned for the 2009-10 school year on a variety of federal-Title, after-school, career & technical education, extended-year, family-support, pre-K, and strategic-compensation programs.[3]

(See footnote for a hyperlink to view the full list of programs.)

District-wide surveys of students, parents, and staff help to guide the prioritization of program evaluations. These surveys include the annual AISD Student Climate Survey,

AISD Parent Survey, AISD Staff Climate Survey, AISD High School Exit Survey, and AISD Central Office Work Environment Survey. They are used "to inform district staff regarding perceptions of the school environment and customer service on each campus, and to examine the work environment of central office departments."[4]

Hanover contacted the DPE to request an interview with the director to find out more information about the department's practices. In place of a phone interview due to time constraints, the DPE responded to a questionnaire via e-mail. The DPE confirmed that AISD commonly evaluates federal-, state-, and locally-funded programs as well as some programs or initiatives supported by private funding (e.g., Gates, Dell). Most evaluations are formalized, while others such as data requests are of an ad hoc nature.

Concerning the prioritization of program evaluations, the superintendent and cabinet-level executive staff annually identify major district initiatives and programs supporting these initiatives, guided by the district's strategic plan and the district's improvement plan. Programs to be evaluated are prioritized based on need for evaluations (e.g., required by law or school board policy), availability of resources in budget for the school year to support evaluation, and availability of data for effectively evaluating the program/initiative.

After determining prioritization and feasibility of evaluation, an evaluation plan is developed in collaboration among program/requesting staff or department, program evaluation department staff, and the chief performance officer. In the process, the evaluation staffing and other resources are made clear, the products or "deliverables" are determined (e.g., measures, reports, etc.), and the specific timeline for the evaluation is agreed upon. Evaluation plan reviewed/approved by the superintendent and cabinet-level executive staff to ensure it meets district needs.

Program evaluations are accomplished by specific individuals or teams of staff, depending on the scope of work and funding available for the evaluation. In the former scenario, an evaluation is supervised by one person and carried out by several team members. Otherwise a team of staff collaborate across the evaluation, with each person taking on a portion of the evaluation activities. In either case, the evaluation process seeks to answer how program goals are tied to the district's strategic plan; how objectives—with measurable outcomes—relate to those goals; and if the programs are implemented with fidelity.

The answers to these questions are found through the collection of data through district and campus sources. Whether for formative or summative evaluations, evaluators have access to large student databases, a human resource database for staff information, and financial data. Surveys are a common data collection tool, either district-wide as described above or as smaller program-specific surveys. When

resources and time are available, an evaluation plan may involve other tools like focus groups, interviews, site visits, observations, and other qualitative data collection methods.

As for evaluation timelines and budgets, most evaluations are conducted on an annual cycle, and the department's evaluation budget is comprised of local, state, federal, and some private-funded monies that can fluctuate from year to year. Each evaluation project has different amounts of funding for staff and evaluation support, thus there is no set amount of staff and other funding for all evaluations. Ad hoc requests are typically narrower in scope and are done in a much shorter timeframe. Other projects may have less strict reporting deadlines, especially if there are ongoing formative reports going to program managers throughout the year.

Once completed, evaluation reports are made available in hard-copy as well as electronic copy (via the DPE website). Regular updates are provided to the superintendent, cabinet staff, and board members. Formative, ongoing reports are provided to program managers during the year, and summative year-end reports provided to program managers. Some evaluation reports are submitted to funding agencies (e.g., state, federal, other) as required, and others may be presented at professional conferences by staff or even submitted for publishing in journals.

1 Austin Independent School District (AISD). "Program Evaluation." http://www.austinisd.org/inside/accountability/evaluation/index.phtml
2 AISD Department of Program Evaluation (DPE). "Plan of Work: 2009-2010." p. ii. http://www.austinisd.org/inside/docs/ope_evaluation_plan_09_10_20091027.pdf
3 Ibid, p. 1
4 AISD DPE, Op. cit., p. 40

MADISON METROPOLITAN SCHOOL DISTRICT

Appendix C - Draft MMSD Curricular Review and Renewal Cycle

MADISON METROPOLITAN SCHOOL DISTRICT
Program Evaluation and Curriculum Review Cycle
Activities & Timelines

In addition to the cyclical tasks described below, the proposed process includes tasks that are performed annually for literacy and math. We specifically recommend that the value added analysis be conducted each year in both reading and math. Further, we recommend that an instructional practices survey be conducted and analyzed in coordination with the value added analysis. This task would be performed to provide insights into why value added data varied across schools and classrooms. If schools or classrooms are performing better, on average, than others with respect to growth in student learning we must determine if this is systematically related to specific instructional practices. This task is included in the Resource/Capacity Action Items within the MMSD Strategic Plan.

Year	Task	Activity
1		**Research & Evaluation**
	1	Initiate program evaluation per MMSD Board of Education approved process
	2	Confirm evaluation questions to be pursued with Board of Education
	3	Define measurement approach in conjunction with Program Evaluation Advisory Committee and curriculum-specific teacher leadership team
	4	Allocate resources to support evaluation team work including determination of what external third party resources might be used for evaluation tasks, if any
	5	Implement data collection in support of the measurement plan
	6	Analyze data and generate summary of findings
	7	Review draft among Evaluation Advisory Committee and curriculum-specific teacher leadership team
	8	Provide final for Board of Education
		Curriculum & Assessment
	1	Collaborate with Assistant Superintendents to integrate program evaluation with the Strategic Plan, District, School and Department Improvement Plans
	2	Collaborate with Research & Evaluation to conduct evaluation
	3	Collaborate with central office and schools to cycle teacher leadership work with major curricular initiatives (e.g., new course proposals)
	4	Allocate resources to support curriculum-specific teacher leadership team work
	5	Establish teacher leadership teams inclusive of multiple perspectives
	6	Establish Advisory Team inclusive of multiple perspectives (administrator, family, community, higher education, student)
	7	Review state and local assessment data to determine patterns and trends across schools and student subgroups
	8	Review local, state, and national curricular standards
	9	In conjunction with Research & Evaluation, conduct a teacher instructional practices survey and review results
	10	Conduct secondary research of instructional strategies and validate the quality of that research

Year	Task	Activity
	11	Review program mission and program goals
	12	Review scope and sequence of courses
	13	Identify gaps and redundancies of program interventions and instructional practices
	14	Prepare draft documents, determine resource needs/budget amount
	15	Present draft documents and resource request to the Board of Education as a discussion item
2		**Research & Evaluation**
	1	Refine evaluation design and data collection based on the review of Year 1 findings, review by the Program Evaluation Advisory Committee and curriculum-specific teacher leadership team, and direction of the Board of Education
		Curriculum & Assessment
	1	Complete curriculum revisions and determine resources
	2	Review recommendations by the Board of Education with discussion
	3	Determine resource needs and select vendors
	4	Narrow down to vendors for field testing, schedule presentations
	5	Field test resources (as appropriate)
	6	Coordinate with Purchasing department for negotiating/pricing
	7	Coordinate with Technical Services Division for specifications and compatibility on digital curricular resources
	8	Determine physical facility needs
	9	Make final recommendations – May
	10	Prepare curriculum documents and resource – May
3		**Research & Evaluation**
	1	Provide on-going support in secondary research tasks and conduct evaluation of specific interventions that might be deployed in a field- or pilot-test mode
		Curriculum & Assessment
	1	Finalize vendor contract – purchasing department
	2	Preview draft with Deputy Superintendent/Chief Learning Officer
	3	Finalize curriculum document and resource adoption
	4	Present adoption to district leadership groups
	5	Present documents and resources to the Board for September/October approval
	6	Order curriculum and assessment resources
	7	Plan spring/summer professional development
	8	Distribute new resources to the buildings before summer break
	9	In-service staff on new adoption
4		**Curriculum & Assessment**
	1	Continue implementation of adopted curriculum and resources
	2	Plan and implement professional development for new staff
	3	Monitor curriculum and make adjustments
	4	Develop curriculum maps and course syllabus, publish for staff
5		**Curriculum & Assessment**
	1	Continue implementation of adopted curriculum and resources
	2	Continue professional development for new staff
	3	Monitor curriculum and make adjustments
6		**Curriculum & Assessment**
	1	Continue implementation of adopted curriculum and resources
	2	Continue professional development as needed
	3	Prepare for new review and evaluation cycle to begin in June of Year 1

MADISON METROPOLITAN SCHOOL DISTRICT

Program Evaluation and Curriculum Review Cycle

Program Evaluation	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016
Year 1	Literacy K-12	Science K-12	World Language	Social Studies K-12	Career & Technical	Physical Education & Health
Year 2		Literacy K-12	Science K-12	World Languages	Social Studies K-12	Career & Technical
Year 3			Literacy K-12	Science K-12	World Languages	Social Studies K-12
Year 4				Literacy K-12	Science K-12	World Languages
Year 5					Literacy K-12	Science K-12
Year 6						Literacy K-12
Program Evaluation	**2016-2017**	**2017-2018**	**2019-2020**	**2020-2021**	**2021-2022**	**2022-2023**
	K-12 Mathematics	K-12 Fine Arts	K-12 Literacy	K-12 Science	World Language	Career & Technical Education

Appendix D - Summary for Program Evaluation

Program Evaluation Protocol
November 2008

Why do we evaluate?

There are at least two purposes for conducting program evaluations within K-12 school districts – formative and summative.

- Formative Evaluation - Formative evaluations provide feedback on the development and implementation of interventions or programs. This is particularly useful during a start up phase. The methods used in formative evaluations are often qualitative in nature including focus groups, interviews, observations, and surveys. The information collected from a formative evaluation is designed to provide a continuous process improvement feedback loop.

- Summative Evaluation – This second general approach to program evaluation is designed to determine the effects of some intervention or program on an outcome, particularly around student achievement (e.g., test scores, GPA, course completion, graduation, etc.) Other student outcomes that might be used as the summative measures are student attendance or behavior.

What do we evaluate?

The single most important step prior to any evaluation study is to make clear the goals and objectives of the evaluation. This is in part driven by the goals and objectives of the intervention itself, but also is largely defined by the actions or decisions which are desired as a result of the evaluation. Do we intend to make decisions around what curriculum should be used? About what professional development investments have the greatest effects on student achievement? Where to allocate resources and positions to have the greatest effects? To determine how to adjust a program to make it better? Establishing the purpose is a critical first step. There are at least three categories of interventions which can be evaluated: core practices, specific targeted, and professional development.

A variety of interventions might be defined for study within a program evaluation. These could include both core practices such as school size or organization (e.g., four block schedule), class size (e.g., 15 to 1 full day, 15 to 1 block/half day), a curriculum (e.g., Connected Math Program), or an instructional strategy (e.g., SERP homework study). There are other more specific targeted interventions that could be evaluated for their effects on student outcomes such as a literacy intervention (e.g., Reading Recovery or Read 180), a engagement intervention (e.g., AVID), an instructional support intervention (e.g., Extended Learning Summer School, after school tutoring)

Another class of interventions which can be studied is those focused on developing the capacity of teachers and instructional support staff. It is possible to determine the effects of a professional development intervention using a carefully designed study. The recent use of value-added growth analyses provides such a tool. Key to this is gathering systematic information on which staff received which professional development experiences and the degree to which this knowledge and skill was then deployed with the desired fidelity in schools and classrooms. There are several examples of professional develop initiatives which could be cited in this group: primary grades Balanced Literacy, the Above the Line behavior effort; and the role of the elementary Instructional Resource Teacher (IRT) and the middle school Learning Coordinator.

There are other types of research questions that are not clearly categorized as evaluation research which merit investigation. Examples include the predictive analyses we have conducted on which students are at risk of not reading by the end of grade 1, are at risk of not completing graduation requirements on time,

etc. The graduation/dropout classification study helped to define who drops and why so interventions can be developed and targeted at specific types of students. There are many other district and school improvement related investigations that have been and should be conducted which are not true program evaluations. These topics must be considered in planning an overall district research agenda.

How do we evaluate?

There are a variety of methods that can be employed to evaluate the effectiveness of programs and interventions. As was stated earlier under the evaluation purpose discussion, the reason for the evaluation can often drive the method of evaluating. Clearly, for formative program development questions feedback is essential on how things are working as systems are being implemented. Focus groups, informal observations, surveys, and interviews are leading candidates for collecting such information. Evaluating trends across the data can be done by analyzing common themes, opinions, and even word choice.

Summative evaluations tend to utilize quantitative methods of analysis and more rigorous designs. They focus on determining what quantitative effect the interventions had most often on an achievement outcome, but also on other student outcomes, too. The analyses range from simple correlational studies, e.g., participation in intervention related to proficiency performance level on the WKCE test, to the much more complex value-added growth models we have recently employed.

Another important factor in determining summative program evaluation effects is in being able to isolate the effects to students who were exposed to a treatment relative to those who were not. The "gold standard" of such evaluation research designs is the randomized trial such as those used in pharmaceutical and other medicals studies. In such cases students would be randomly assigned to a treatment or control group. This is often very difficult to do for all sorts of practical reasons. One method of obtaining this type of assignment is to use two groups where one group receives the treatment at time A and the second group receives it at the later time B. No one is denied services, they simply receive it at two different points in time. The initial group serves as the treatment group while the latter group becomes their comparison group.

Quasi-experimental designs are another method that tries to approximate randomized trials by simply taking two groups and matches them in one manner or another such that they are statistically similar. Randomization does not occur so the potential exists that student self-selection may be biasing any difference in results.

Another important form of research we must always engage in is secondary research, often in the form of a literature review. A careful analysis of existing research can provide insights into what works and how we might deploy an intervention within the MMSD. A quantitative approach known as meta-analysis might even be applied to existing studies to determine which interventions provided greatest impacts. Meta-analysis compiles the statistical outcome data from previous studies and ascertains average effects across the studies included in the pool. Often, such meta-analyses already exist in the secondary research. However, regardless of approach, the most important factor to consider in secondary research review is being able to evaluate the quality of the previous studies for rigor. Poorly conducted studies even in large numbers cannot yield valuable information upon which to make decisions regarding interventions.

Do we evaluate alone?

The simple answer is, "No." The MMSD research team is very limited is personnel. Wherever possible collaboration is desired. Fortunately, we have several precedents and opportunities for collaboration. These include:

- Wisconsin Center for Education Research (WCER)

- Minority Student Achievement Network (MSAN)
- Other university faculty and student-sponsored research requests through the MMSD External Research Committee (ERC)

Another very exciting opportunity is the development of the Wisconsin Department of Public Instruction's (DPI) longitudinal data system (LDS) data warehouse research database. Funded by a grant form the US Department of Education, the LDS will provide a rich set of student data – all protected within confidentiality rules – that can be the basis for evaluating the effects of interventions across districts in Wisconsin. This arrangement unleashes the potential for greater Wisconsin Idea efforts which involve the K-12 districts, DPI, and the University of Wisconsin in studying what works.

How do we decide what to evaluate?

A district program evaluation plan for any school year must be driven by the district strategic goals and objectives. The strategic plan provides the framework for determining which projects are priorities. Yet there are other factors to consider in determining what to evaluate.

The volume of evaluations which can be performed is a function of the resources available and required to complete them. The form of evaluation is often driven by resource availability. Less complex evaluation methods are less expensive to conduct. More complex evaluations often require the use of outside consulting resources. Our current practice, in place for over ten years, is to conduct between one and three evaluations per school year.

Evaluation decisions should be driven by purpose. The Board of Education certainly determines many such decisions. The administration can also recommend evaluations. The Management Team and Instructional Council are the appropriate vehicles for developing administration recommendations for program evaluations.

In general, it is important that these stages of evaluation be followed prior to implementing a district-wide intervention:

1. Secondary research – literature review, focus on meta analysis and USDOE What Works Clearinghouse

2. Field test – formative evaluation to determine how best to deploy

3. Pilot test – additional formative feedback and a quantitative summative evaluation using a limited number of schools/program sites and students in at least a quasi-experimental design, prferravly a randomized assignment design

4. Full deployment – formative feedback for collecting fidelity of implementation data and a time-lag quasi-experimental design quantitative outcomes analysis