

Running head: AN ILLUSTRATIVE CASE STUDY

***An Illustrative Case Study of the Heuristic Practices of a High-Performing
Research Department: Toward Building a Model Applicable in the Context of
Large Urban Districts***

Marco A. Munoz and Robert J. Rodosky

Jefferson County Public Schools

Paper presented at the annual meeting of the Consortium for Research in
Educational Accountability and Teacher Evaluation (CREATE) 2011 National
Evaluation Institute.

Abstract

This case study provides an illustration of the heuristic practices of a high-performing research department, which in turn, will help build much needed models applicable in the context of large urban districts. This case study examines the accountability, planning, evaluation, testing, and research functions of a research department in a large urban school system. The mission, structural organization, and processes of research and evaluation are discussed in light of current demands in the educational arena. The case study shows how the research department receives requests for data, research, and evaluation from inside and outside of the educational system, fulfilling its mission to serve the informational needs of different stakeholders (local, state, federal). Four themes related to a school district research department are discussed: (1) basic contextualization, (2) deliverables of work, (3) structures and processes, and (4) concluding reflections about implications for policy, theory, and practice. Topics include the need for having an evaluation model and the importance of having professional standards that guarantees the trustworthiness of data, research, and evaluation information. The multiple roles and functions associated with supplying data for educational decision making is highlighted.

Keywords: Data Analysis; Decision Making; Educational Planning; Educational Research; Evaluation; Evaluation Methods; Formative Evaluation; Management Systems; Program Evaluation; Public Schools; School Districts; Student Evaluation; Summative Evaluation; Urban Schools

***An Illustrative Case Study of the Heuristic Practices of a High-Performing
Research Department: Toward Building a Model Applicable in the Context of
Large Urban Districts***

This case study is about a research department in the large urban school system. Accountability is understood as how schools, teachers, parents, central office administrators, and the community must be held responsible for the education of the district's children. The research department has structures and processes to support ongoing evaluation for accountability, program improvement, decision-making, and meeting external and internal mandates. This case study, a follow-up and extension of a prior study (Rodosky & Muñoz, 2009), concludes with a discussion of the essential practices of a research department in a large urban school system.

Methodological Approach

For this active-participant case study, the qualitative research design permitted to explore the interviewee's opinions in depth and allowed for elaboration of existing concepts. This inductive method allowed gaining a broader perspective and a deeper understanding of the phenomenon. Interviews and document analyses served as the primary source of data for the study. The interviews followed a semi-structured format (Merriam, 1988) to encourage genuine and unhindered responses. The semi-structured format used during the interviews involved the interviewer asking questions to direct

responses toward the topic of interest, but did not involve a non-flexible protocol. Patton (2002) notes the following:

The data for qualitative analysis typically come from fieldwork. During fieldwork, the researchers spends time in the setting under study—a program, an organization, a community, or wherever situations of importance to a study can be observed, people interviewed, and documents analyzed (p.4).

The grounded-theory paradigm (Glaser & Strauss, 1967) was used as a guiding framework. Under this paradigm, research is guided by initial concepts, but can shift or discard these concepts as data are collected and analyzed (Marshall & Rossman, 1989). Data collection and analysis occurred simultaneously. This process was continued throughout the study. In employing the constant comparison method (Glaser & Strauss, 1967) for comparing segments within and across categories, the meaning of each category, and distinctions between categories were studied in deciding which categories were most important to the study.

Coding processes included identifying concepts embedded within the data, organizing discrete concepts into categories, and linking them into broad, explanatory themes (Strauss, 1987). Content analysis allows the researcher to identify patterns or themes that emerge from the data (Patton, 2002). Emerging categories served as the filtering lenses through which the interview transcripts, filed notes, and documents were examined. Over time, the number of coding

categories was reduced by eliminating and merging categories and by clustering still other categories based on perceived connections. This repetitive process eventually led to the construction of qualitatively distinct themes.

Case Study Context

The school district is located in Kentucky and is the 31st largest district in the nation, comprised of an urban core, suburban housing, and diminishing rural areas. It has 150 schools serving approximately 97,915 students, with an annual budget of over \$1,000,000,000. It has about 13,700 full-time and 5,800 part-time employees.

The district has a student assignment plan based on “managed choice,” a plan that facilitates the racial desegregation of its schools by providing students with transportation from their neighborhood homes to other parts of the district. This plan has been the focus of extensive court supervision up to the present time. School-Based Decision Making (SBDM) is part of the Kentucky Education Reform Act (KERA) of 1990, and individual SBDM teams set school policy consistent with district board policy, while district officials can suggest academic programs and interventions: Individual schools, through SBDM, have ultimate control over the adoption of curricula and programs in their building. There is an elected Board of Education, a superintendent and one assistant superintendent for each school level. Accountability, Research, and Planning is one of ten departments that report directly to the superintendent. The Department has 20 employees focused on providing reliable, valid, and useful

information to decision-makers in a timely manner. This is our organizational mission and it is accomplished on an everyday basis.

The organizational, structural location of the Accountability, Research, and Planning Department is critical to its effectiveness. It is crucial to have direct-line reporting to the superintendent, particularly when everyone comes to us for data requests from inside and outside the schools. To be effective, the top leadership of the Department must have authority to carry out these complex tasks and a direct line to the superintendent. In our data-based, accountability-focused world, having such authority is crucial. Our customers are inside and outside the schools, including community-based organizations. For example, we have developed partnerships with over 70 community groups who use our online data system to tutor, counsel and mentor students.

Student academic achievement is the schools' primary purpose. To monitor and assess this purpose, the district tracks and reports academic progress regularly through multiple reports, e.g., District Report Card, School Reports Cards, No Child Left Behind (NCLB, 2001), Adequate Yearly Progress (AYP) results, and an online database of student formative assessments.

From a data use and continuous improvement perspective, the assumption is that diagnostic, formative assessment is as equally valuable as summative assessment results. As a result, the district pursues a balanced assessment system that includes multiple types of tests and assessments. According to Stiggins (2006), a balanced assessment is an integration of

classroom assessment, interim benchmark assessment, and accountability tests into a unified process that benefits student learning. A balanced assessment system provides for the information needs of assessment users at all of these levels.

Internal Organization and External Influences

The Department includes Research, Planning, and Accountability. The Research Unit does institutional research and data warehousing activities and promotes district internal and external research and evaluation activities by providing valid and reliable data efficiently and in a timely manner in an atmosphere that is inviting, receptive, and responsive to the data needs of its customers in the schools and community. It also designs, administers, and reports surveys providing feedback for planning and evaluating programs to the Board of Education and local schools. These annual surveys are in the areas of quality of education, school safety and staff job satisfaction. This unit conducts its own research activities and acts as a clearinghouse for external research initiatives, providing initial screening and support to a variety of research requests.

The Planning Unit which coordinates state-required school and district plans, and provides service to the district for its Southern Association of Colleges and Schools (SACS) accreditation. The Planning Unit coordinates the format, timelines, quality reviews, and training for the development of the Comprehensive School Improvement Plan (CSIP). It also compiles the

Comprehensive District Improvement Plan (CDIP), which outlines proposed work improvements in the core content areas of reading, writing, and mathematics. The CDIP also lays out a district plan to provide resources to its most struggling schools, while also coordinating the district's dialogue and coaching process on priority schools (i.e., low-performing schools) identified through assessment data. Finally, the unit provides grant writing technical support and evaluation services to numerous district grants and programs which are use for program improvement.

Accountability (testing) is another important structure in the Department. The Testing Unit focuses on the coordination, implementation, and management of logistics associated with the various state and district assessments. We use both statewide academic and non-academic assessments, and a district-based continuous assessment process, along with several other assessment programs, all with their own rules, instruments, and timing.

Federal mandates influencing our work include the Elementary and Secondary Education Act (ESEA) Title I program and No Child Left Behind Act (NCLB). The latter requires specific evaluation of Annual Measurable Objectives (AMOs) — a task that has created much work. At the state level, KERA (1990) requires regular district assessments of student academic and non-academic performance. Other laws, external requirements, and data demands drive our work and given the uniqueness of our district, this makes meeting these often competing demands difficult.

The Deliverables of the Research and Evaluation Work

Our district and department have systems and processes in place to respond to accountability demands. Our work is guided by our mission to facilitate data-based decision-making. We use primarily Stufflebeam's Context-Input-Process-Product (CIPP) Model (Stufflebeam, 1983; 1985; 2001; 2002; 2004; 2005; Stufflebeam & Shinkfield, 2007; Stufflebeam et al., 1971) to provide users with useful, valid, and reliable data in a timely manner.

Consistent with its prospective, improvement focus, the CIPP Model places priority on guiding planning of enhancement efforts. In the model's formative role, context, input, process, and product evaluations respectively ask: (a) what needs to be done? (b) how should it be done? (c) is it being done? and, (d) is it succeeding? Prior to and during the decision-making and implementation process, the evaluator submits reports addressing these questions to help guide and strengthen decision making, keep stakeholders informed about findings, and help staff work toward achieving a successful outcome.

The model's intent is to supply evaluation users—such as policy boards, administrators, and project staffs—with timely, valid, reliable information of use in (a) identifying an appropriate area for development; (b) formulating SMART goals, activity plans, and budgets; (c) successfully carrying out and, as needed, improving work plans; (d) strengthening existing programs or services; (e) periodically deciding whether and, if so, how to replicate or expand an effort; and, (f) meeting a financial sponsor's accountability requirements.

The model also advocates and provides direction for conducting retrospective, summative evaluations that serve a broad range of stakeholders. They include, among others, (a) funding organizations, (b) persons receiving the sponsored services, and (c) policy group, and researchers outside the program being evaluated. In the summative report, the evaluator refers to the store of formative context, input, process, and product information. The evaluator uses this information to address the following retrospective questions: (a) was the program keyed to clear goals based on assessed beneficiary needs? (b) was the effort guided by a defensible procedural design, functional staffing plan, effective and appropriate process of stakeholder involvement, and a sufficient, appropriate budget? (c) were the plans executed competently and efficiently and modified as needed? and, (d) did the effort succeed, in what ways and to what extent, and why or why not?

Our department produces multiple reports for the superintendent, administrators, and the Board of Education. For the superintendent, each evaluation report is composed of three sections: an executive summary, a managerial report, and a technical report. The executive summary concisely describes key elements of the full evaluation report: (1) background information, (2) evaluation questions, (3) evaluation method/instrumentation, (4) evaluation findings, and (5) evaluation recommendations. The managerial report provides more detailed information focused more on processes than outcomes; it is written for a non-technical audience. The third is the technical report, which

provides detailed statistics. These are often presented at national conferences and/or published in peer-reviewed journals.

Key deliverables in our district are two locally produced books, School Profiles and Data Books with multiple academic and non-academic data aggregated at the school level. These summarize the comprehensive demographic, cognitive, and non-cognitive characteristics and achievement of each school and the district as a whole.

Comprehensive School Surveys are another basic deliverable. We develop, distribute, administer, collect, analyze, and report on feedback surveys from various stakeholders (i.e., students, parents, teachers, and staff) about school and district practices. Areas covered by these surveys include quality of education, school climate, job satisfaction, and safety and these are used to make informed decision to improve educational services.

Our latest deliverable is a formative assessment system using new technologies to produce results. For example, teachers can produce individual student test answer sheets aligned with district-made formative tests and web-based “dashboard” reports. All these efforts keep our curriculum (standards), instruction, and assessment domains as coordinated systems.

Data for Decision-Making and Planning

The superintendent and Board of Education need data for policymaking. Some come directly from the state Department of Education, but we generate, analyze, and report most of what they want and use. We develop our teacher

decision-making data. Since most summative reports are not useful for classroom teachers deciding on instructional steps for individual students, we developed a system of formative assessments with a variety of different measures that generate school-, class-, and student-level data for use by teachers, counselors, and principals.

In general, data from state and district assessments are moved electronically into the district's system and are readily available to teachers for planning instruction and making curricular decisions. Student data from multiple assessments are individually entered into the system (either by keyboard entry or scanning) and available for teacher use. All of these timely, easily accessible data allow district leadership, principals, staff members, parents, community members, and SBDM council to analyze student performance in an individual school, determine progress toward CSIP goals and benchmarks, and evaluate the effectiveness of the district's or school's interventions and programs.

District organizational improvement is another focus. We help answer questions about the mission of the schools and our department and what must be done to remain competitive. This is more and more crucial because public schools are no longer a monopoly, competing with private and parochial schools for students (i.e., market share needs to be constantly monitored and establishing magnet school is a must). Basic to organizational development is the identification and assessment of best practices across the country. We help decision-makers with data on these and help also with long-range planning,

providing reports about finances, operations, and the external environment. Additionally, we help school- and community-based planning by collecting, analyzing, and using data.

Evaluation and Research

Evaluation and research are core components of our everyday work. Using the Joint Committee Standards for Program Evaluation (1981; 1994; 2011) as guide, we go about evaluating key initiatives and programs to ensure that we are helping our students. As mentioned earlier, we use the CIPP model (Stufflebeam, 1983; 1985; 2001; 2002; 2004; 2005; Stufflebeam & Shinkfield, 2007; Stufflebeam et al., 1971):

- Context evaluations assess needs and opportunities to help decision makers define objectives and priorities.
- Input evaluations assess alternative approaches and strategies for their feasibility and cost-effectiveness to meet identified needs. Decision-makers use input evaluations in choosing among competing plans, writing funding proposals, allocating resources, and organizing work.
- Process evaluations assess the implementation of plans to help staff carry out activities and help stakeholders improve program performance.
- Product evaluations identify and assess outcomes both intended and unintended.

We have an evaluation team to develop structures and processes to provide useful information to support student learning. It establishes results-

oriented learning processes, and data are used for practical improvement of schools and classrooms. The evaluation team develops a program profile (including the program's goals), conducts evaluation processes and activities, and reports evaluation's results and recommendations to key stakeholders. During the evaluation process, evaluators visit all schools involved and observe classrooms, and interview students and teachers. During that process, academic data are used to assess progress toward meeting improvement in student learning targets; we want to avoid "black box" evaluations that are limited with process data (Muñoz, 2005). Efforts associated with backward planning as well as statistical and practical significance are incorporated in the evaluations (Muñoz, 2005). The evaluation team makes recommendations about the future of each program, including (a) reduction, (b) expansion, (c) elimination, (d) adjustment, and/or (e) continuation. Clearly, evaluation is central to our department's work on accountability. This work is linked with communication efforts about what we do, how we do it, and what we have learned. Our message has a purpose and an audience – promoting decision-making based on valid, reliable, and timely data.

We use the CIPP model in coordination with professional standards. Professional standards for evaluations are principles commonly agreed to by qualified specialists in the conduct and use of evaluations for the measure of an evaluation's (a) utility, (b) feasibility, (c) propriety, (d) accuracy, and (e) evaluation accountability. Basically, the CIPP model is an organized approach to

meeting the evaluation profession's standards as defined by the Joint Committee on Standards for Educational Evaluation (1981, 1994, 2011). The model calls for evaluators to meet the professional standards of evaluation and subject their evaluations to both formative and summative metaevaluations.

As feasible, the evaluation should be subjected to an external, independent metaevaluation. This helps avoid any appearance or fact of the evaluator's possible conflict of interest having influenced the content of the external metaevaluation report. The external metaevaluator's report should be made available to all members of the right-to-know audience. A checklist designed to facilitate systematic application of the Joint Committee *Program Evaluation Standards* in metaevaluations is available (www.wmich.edu/evalctr).

Structure and Processes

Looked at as a process of development, strong leadership is critical in the early phases of developing new structures and processes. This process is guided by a vision that comes from our colleagues at the top of the organization as to how to respond to the changing environment of public education. The most fundamental aspect of day-to-day managing for accountability is that you must know for whom you work. In order to carry out this mission, the superintendent gives the leadership in our department the full authority to carry out responsibilities. Within this framework, we strive always to provide a quality product within a specified timeline and budget. We have structured and manage

the daily operations of the department is driven by a focus on outcomes that benefit kids.

The basic structure of the Accountability, Research, and Planning Department is organized around outcomes. Our bottom line is customer service focused on student learning. The department's delivery system operates on well-defined objectives organized within a framework that is both vertically and horizontally aligned. The vertical alignment comes directly from our superintendent; the chain of command from the Superintendent to us is clearly established and concerned with the flow of authority and responsibility. Horizontal alignment arises from us knowing the strengths and weaknesses of our departmental team.

A variety of structures and processes help us deliver accountability data, analyses, and reports related to our accountability, research, and planning functions. These processes require constant monitoring to ensure timely delivery. If we are doing the same work in the same way, over a period of several years, we are probably inefficient and insufficient. We set objectives for all structures and processes. These are specific, feasible, measurable, and time-bound.

We include key stakeholders in these need assessment, planning, implementation, and outcomes discussions: this meets the very important Standards for Program Evaluation requirement of stakeholder involvement. It also gives us an opportunity to create buy-in with key staff and stakeholders and

objectives that are truly based on stakeholders' needs and wants. We also keep a sharp focus on the needs of our customers and continually reassess whether discrepancies exist between what is and what is desired. The objectives of all our structures and processes must satisfy the needs of our key customer base — our students!

Staffing and Budgeting

We hire only the best people, and our staff can analyze data with sophisticated statistical techniques and also develop data systems easily accessed on the web. More importantly, staff are trustworthy. They can represent the department in meetings, make decisions about the technical aspects of their work, and meet the other requirements of evaluation research without being micromanaged. Staff are always up to date with professional and organizational knowledge and skills, including new technology, our very high quality evaluation work is seen in professional presentations and publications. Staff are critical to promoting improved services to students and schools. Our leadership constantly ask staff, "What do you need to get the work done in a quality and timely way?" The leadership supports new developments if they make sense and will likely provide better service to our customers. "Mental space" is provided for staff to envision and implement alternative organizational structures and processes. There is a general bias toward action, through trial and error, and always with our kids in mind. We are productive simply because we

have a competent and hard-working staff. Managing day-to-day means allowing and supporting staff to do their best!

The work precedes the budget. We need to work smarter, rather than harder (Drucker, 2006). This extends to how we use funds. In any given year, we typically receive the same amount of funding as in the previous year. We use flexibility in crossing accounting codes to avoid asking for extra funding. Why? Because in school districts, the more you ask for, the more you will become a target for cost reductions. We need to constantly rethink what is necessary to better respond to new needs within budget constraints. That is the true test of the efficiency and productivity of our system — not whether our budget is larger. We budget the cost of processes associated with the work, including data acquisition, input, storage, analysis, and reporting. If truly necessary, we seek additional funds to make specific studies happen. Why? Because staff need to have what is needed to do their work.

Several strategies are used to bring in needed additional funds, directly or in-kind. One way is by creating partnerships with schools and other departments through joint projects. Our operational rule is that schools or departments need to match (at some level, usually about 10% or so) our financial support. Our budget typically includes funding for professional development activities and for technology, both basic support elements for our work in planning, research, evaluation, and accountability. These line items, which are targets for cuts in some organizations, are the last to be considered when we have to give up

funding, for the reasons discussed above. Without top staff and good, appropriate technology, we simply could not do our work.

The Core Beliefs that Guide Research and Evaluation Work

One needs core beliefs and values to guide one's everyday work. We believe that our department must be morally accountable for the important "data" work. Our internal customers are the superintendent and the assistant superintendents, board members, executive directors, and program coordinators. But our most important internal customers are students, teachers, and parents. Our external customers are the local, state, and federal leaders interested in developing our public schools. This is department policy, and this philosophy is found in all the ways in which we respond to all of our customers. We believe in data-based decision-making. To that end, we continuously work to bring empirical data to all relevant work conversations. We resist talking in terms of perception-based hunches. We slay myths with data!

We also believe that the ad hoc work, if proven, should become business as usual. One example was the development of the school profiles and data books. Several years ago, we would receive several requests from community members, parents, principals, and other school staff for the same information. We had to respond to each of those requests, an inefficient use of our time and resources. As a result, we built a process to generate a consistent response (i.e., institutionalizing routines). We now produce these books early in the school year. In this way, we created a needed, useful product and turned the ad hoc

into a regular and recurring process. We also believe in recruiting and hiring the best people. The right individual in an evaluation unit means someone who is capable of analyzing and improving his or her own work (i.e., meta-evaluation). In this sense, it is important to stay up-to-date with major issues associated with accountability, assessment, and evaluation in the schools and the community. As much as possible, we try to accomplish staying up-to-date through regular involvement in professional organizations such as the American Educational Research Association (AERA), the American Evaluation Association (AEA), and the Consortium for Research in Educational Accountability and Teacher Evaluation (CREATE).

The essential and best practices involve multiple elements: (a) a passion for kids as a moral grounding; (b) credibility when presenting results (both good and bad); (c) a trifocal perspective (i.e., the tree, the forest, and the interconnectedness of the two); (d) a polychronic structure (i.e., multiple clocks all keeping the school calendar in mind); (e) a transparent, democratic process (i.e., eliminate gate keeping), (f) data-driven policy decision-making; and, (g) self-reflection while comparing one's self constantly against best practices. A passion for kids is naturally the core element here. Evaluation for accountability in the school setting means helping kids. If you don't like kids, we are in the wrong business!

Credibility is another important element because it is the currency of the work. We do honest work and can defend our work. We do not edit or in any

way distort data. Another core element is tri-focal lenses to view our work.

When examining an issue, we make sure to look from both concrete/unique and abstract/general perspectives. We also must see the synergy and inter-relations between the individual and the systemic.

We also bring a polychronic perspective to decision-making and organizational processes; there are multiple “watches” ticking in our mind when it comes down to decision-making. For example, we build our work around the school calendar. We know that there are some decisions that need to be made in March (e.g., planning, funding), some during the summer months, some at the beginning of the school year, and some at the semester. This is unavoidable in a large school system when one is evaluating for accountability.

We like to eliminate gate-keeping through the creation of a transparent, democratic process that facilitates access to data; this access to data, in turn, facilitates our ultimate goal — sustaining a data-driven decision-making environment. We must definitely be self-reflective because we are accountable for our work. Self-reflection is self-evaluation: our own work is also data-driven!

Future work under this heuristic perspective of the research work might include providing more services to different levels of the school system (e.g., students, teachers, administrators, and community members): (a) consulting and training on data analysis and interpretation; (b) consulting and training on use of formative and summative data; (c) support for grant development and compliance; (d) expand the use of cost analyses into reporting of

school/program accountability studies; (e) consult and training on classroom action research (inside out studies); and, (f) support of central office departmental reviews. There is always room for growth when you have a continuous improvement philosophy.

In summary, at the foundation of the Research Department efforts, this is all about a working out of a simple philosophical base: Kids must always come first! Second comes using data for accountability. We need to produce accurate, meaningful, credible, and useful data. We cannot fall in the trap of believing that our work as evaluators is more important than the teacher's work in the classroom. Our work is a means and support toward an end — student learning. Our job is to support teachers as they work to accomplish the most precious work of all: educating children so they can become contributors to society.

References

- Drucker, P. F. (2006). *Classic Drucker: Essential wisdom of Peter Drucker from the pages of Harvard Business Review*. Boston, MA: Harvard Business Press.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research*. New York: Aldine.
- Joint Committee on Standards for Educational Evaluation. (1981). *The program evaluation standards*. Thousand Oaks, CA: Corwin Press.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards (2rd ed.)*. Thousand Oaks, CA: Corwin Press.
- Joint Committee on Standards for Educational Evaluation. (2011). *The program evaluation standards (3rd ed.)*. Thousand Oaks, CA: Corwin Press.
- Marshall, C., & Rossman, G. B. (1989). *Designing qualitative research*. Newbury Park, CA: Sage.
- Merriam, S. B. (1988). *Case study research in education*. San Francisco: Jossey-Bass.
- Muñoz, M. A. (2005). Backward mapping. In *Encyclopedia of Evaluation* (p. 29). Thousand Oaks, CA: Sage.
- Muñoz, M. A. (2005). Black box. In *Encyclopedia of Evaluation* (pp. 34-35). Thousand Oaks, CA: Sage.
- Muñoz, M. A. (2005). Significance. In *Encyclopedia of Evaluation* (p. 390). Thousand Oaks, CA: Sage.
- No Child Left Behind Act. (2001) Pub. L. No. 107-110.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Rodosky, R. J., & Muñoz, M. A. (2009). Myth slaying and excuse elimination: Managing for accountability by putting kids first. *New Directions for Evaluation*, 121, 43-54.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. New York: Cambridge University Press.

- Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2006). *Classroom Assessment for Student Learning: Doing it right – using it well*. Portland, OR: Educational Testing Service.
- Stufflebeam, D. L. (1983). The CIPP model for program evaluation. In G. F. Madaus, M. Scriven, & D. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluations*. Boston, MA: Kluwer-Nijhoff.
- Stufflebeam, D. L. (1985). Stufflebeam's improvement-oriented evaluation. In D. L. Stufflebeam & A. J. Shinkfield, *Systematic evaluation* (pp. 151–207). Norwell, MA: Kluwer.
- Stufflebeam, D. L. (2001). Evaluation models. *New Directions in Evaluation*, 89. San Francisco, CA: Jossey Bass.
- Stufflebeam, D. L. (2002). *CIPP evaluation model checklist*. www.wmich.edu/evalctr/checklists.
- Stufflebeam, D. L. (2004). The 21st-century CIPP Model: Origins, development, and use. In M. C. Alkin (Ed.), *Evaluation roots*. Thousand Oaks, CA: Sage.
- Stufflebeam, D. L. (2005). CIPP model (context, input, process, product). In S. Mathison (Ed.), *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage.
- Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. (1971). *Educational evaluation and decision making*. Itasca, IL: Peacock.
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.