

# School System Evaluation: A Generative Approach

---

## *School System Evaluation: A Generative Approach*

Pam Adams  
David Townsend

---

**ABSTRACT:** An invitation to explore innovative practices in program and system evaluation in a medium-sized school jurisdiction of approximately 11,000 students has resulted in a model with generative characteristics. This paper will describe the process, and several of the outcomes, when a generative evaluation approach is used to assess the effectiveness of key components of a school system. In addition, it will highlight characteristics that distinguish this type of evaluation from more conventional forms.

---

### **Introduction**

The School Division that is the subject of this study has a complex history of cultural change and challenges over the past decade. This includes events such as the forced amalgamation of three distinct jurisdictions into one, the termination of several senior educational officers, and a strike and lockout that was part of widespread educational labor strife. Because and in spite of the difficulties they faced, educational and political leaders sought to transform their Division from one that claimed to be a system on the leading edge to one where students come first. The evaluation reported in this study was one of many initiatives designed to enhance organizational learning, improve instruction, and develop leadership capacity.

### **Background**

For more than thirty years, program and system evaluation in North America has been profoundly influenced by Scriven's (1967) formative and summative classifications. While those distinctions have proven durable, recent developments challenge several assumptions about the evaluation process, calling into question such fundamental concepts as objectivity, purpose, usefulness, impact on participants, roles of participants, and uses of results. Stake (1975) first proposed that education evaluation be responsive, that is, "orienting more directly to program activities rather than program intents; responding to audience requirements for information; and referring to different value perspectives in reporting program successes or failures" (p. 14). In turn, Guba and Lincoln (1989), Hopkins (1989), and Fetterman (2001) have advocated varying forms of empowerment evaluation. These authors contend that people can discover knowledge and solutions based on their own experiences, and that internal stakeholders are able to establish their own goals, processes, and outcomes. Alternatively, in this approach, external evaluators can provide training, coaching, and assistance in an atmosphere of honesty, trust, and support, contributing positively to the formation of a "dynamic community of transformative learning" (Fetterman, 2001, p. 6). Cousins and Earl (1995) have developed a model of participatory evaluation, which also involves external evaluators working in partnership with practitioners. It is "responsive to local needs, while maintaining sufficient technical rigor so as to satisfy critics" (Cousins & Earl, 1995, p. 9). Posavac and Carey (1997) propose a deliberative democratic approach through which the needs of stakeholders are served, valid information is provided, and alternate viewpoints are acknowledged. Mertens (1999) calls for emancipatory evaluation, which aims to address the needs of those people with least power, so they might be better able to influence their own destiny. According to Fitzpatrick, Sanders, and Worthen (2004), these and similar participant-oriented evaluation approaches "have prompted more acrimonious debate than almost any development in evaluation within the last two decades" (p. 146).

For almost two generations, the dominant model of program evaluation has been relentlessly summative, overly-formal, episodic, and generally ineffective in contributing to desirable change. School and system evaluations---when they occur--- are still conducted by teams comprised primarily of members external to the system, assisted by a lesser number of internal participants. Typically, terms of reference are negotiated in advance, letters of transmission

exchanged, contracts formalized, schedules of classroom or office visits developed, and the ordeal begins. Lesson plans and curriculum documents get dusted off, old wardrobes reappear, suits and briefcases dominate hallways and staffrooms and then, after about ten days, all returns to "normal." Two or three months later, an official report is presented to the school board, even as the evaluation event is passing into memory.

As educators who have engaged in forms of evaluation that attempt to use more responsive and participatory approaches, we were pleased to respond when the School Division superintendent invited us to help plan and conduct an evaluation that was, in his words, "Very important, even if it isn't the most urgent matter we have to attend to." In accordance with empowerment evaluation principles, we entered into a process of negotiation that involved cultural awareness and readiness activities. These included discussions with the school board, the administrators' council, the district administrative team, and teachers' groups. Subsequently, a project steering committee was formed and, under its guidance, the next stages of the process were planned and implemented.

The primary tasks of the steering committee were to approve the research question that would guide the project, the methodology that would be used, the letter of invitation that would lead to the formation of an internal action team, the project timelines, and the project deliverables. The question, debated and refined over a six-week period, was What are the processes, structures, strategies, principles, and values that contribute to the Division's pursuit of high levels of effectiveness?

### **The Evaluation Process at Work**

The action team was comprised of two external researchers, two central office administrators, two district curriculum consultants, one principal, and six vice-principals. Educators from within the Division volunteered to be part of the action team, and the final selection of members was based on criteria approved by the steering committee. Most team members had some experience conducting research; four had completed their Master of Education degrees, and four were at varying stages of completion. They agreed to participate knowing they would have to devote several half-days to training, and at least an equal number to data collection and data analysis activities. The negotiated project plan called for data collection through focus groups (Fern, 2001), surveys (Gay & Airasian, 2003), document analysis (Neuman, 1997), and interviews (Wengraf, 2001).

#### *Focus Groups*

Few members of the action team were familiar with the focus group method. Accordingly, to help develop their expertise, a process was used that included an initial training workshop, followed by guided practice, followed by more focused training, followed by independent practice. Specifically, after the initial training workshop, pairs of team members conducted one focus-group session, brought their experiences and findings back to a second training workshop, and reviewed the process. Within the next two months, team members conducted 14 additional focus groups with a total of 150 participants representing parents, students, administrators, trustees, teachers, and support staff. A third training workshop provided direction for organizing and analyzing focus group data using Neuman's (1997) method of thematic conceptualization, open coding, axial coding, and selective coding. Focus group findings were influential in identifying five critical dimensions of community that formed the structure of the survey instrument.

#### *Surveys*

All action team and steering committee members assisted with the refinement of the survey, which was then distributed and collected within a three-week period. Four hundred and two surveys were returned by representatives of all employer groups in the Division. Survey data were analyzed by one team member and one administrative support staff member using Statistical Package for Social Sciences (SPSS 11.5). The survey analysis yielded compelling data on dimensions of district operations specific to mission and vision, culture, leadership, learning, and organizational structure.

### *Document Analysis*

All action team members participated in a document-analysis training process similar to the one that developed their focus-group facilitation skills. This proved particularly effective when team members faced the challenge of document analysis. The first workshop produced a template for data collection, a method for ensuring inter-rater reliability (similar to the method used for marking achievement test essays), and a strategy for sampling more than 5000 school and district documents. In ensuing workshops, team members systematically organized and analyzed data extracted from the documents.

### *Interviews*

In a workshop devoted to determining what had been learned and what still needed to be answered regarding the research question, action team members created nine open-ended interview questions that were subsequently administered to a stratified random sample of 49 school division employees. Team members analyzed interview data by using, once again, Neuman's (1997) coding protocol. Summaries of interview data were verified through a process of consultation involving selected participants and team members.

### *Reporting Results*

While all members of the action team contributed directly to the writing of a final report, a smaller group evolved as the writing team. Five members worked on the first draft of the report, which was subsequently revised and refined several times. The report, together with an executive summary, was then presented to the superintendent and school board. After a final round of minor revisions, verified by the action team and the steering committee, the report was posted on the Division's website.

## **Reflection on the Process**

The results of this study, gathered over a period of thirteen months, provide persuasive evidence of the ways that schools can move forward purposefully in the pursuit of agreed-upon goals. An understanding of the differences between this and more traditional forms of evaluation became apparent as system leaders responded to findings generated in each of the data collection phases rather than waiting for a summative report to identify needs and authorize action. For example, when survey results showed exceedingly low ratings by one distinct group of employees, the superintendent moved immediately to discover the reasons for their dissatisfaction. That investigation revealed a morale problem so serious that it resulted in the dismissal of a supervisor. As another example of the generative character of this process, action team members became convinced that the survey's diagnostic capabilities could help shed light on the Division's progress towards its goal of developing all schools as learning communities. Consequently, action team members participated in the construction of a learning communities rubric that identified dimensions and characteristics of several types of communities. The document was broadly distributed throughout the Division.

In a related way, document analysis provided action team members with powerful insights into the importance of the role of communication in effective school and system operations. It led directly to some simple changes in document accessibility and content of some school-based documents, while several findings drawn from document analysis were quickly incorporated into the Division's decision-making process. Similarly, when interview data showed conclusively that some staff evaluation practices, methods of communication, and administrative team structures were not as effective as they might have been, district leaders took immediate action to rectify each situation.

The extensive generative evaluation process confirmed that this was a school division doing many things well. For example, students and parents were overwhelmingly satisfied with the quality of leadership and educational services provided by the Division. Achievement Test and Diploma Examination results showed positive trends extending over several years. School principals were illuminated as a strong force in the Division culture, particularly in their commitment to the Division's model of balanced decision-making.

Leaders in the Division were adamant that the evaluation's impact was productive. They reported that it contributed directly to system learning (Senge, 1990), capacity building (Lambert, 1998), shared leadership (Elmore, 2000), improved morale (Barth, 2001), and an increased sense of belonging (Deal & Peterson, 2000). Conversely, it did not disrupt system operations, and it did not heighten the distress of evaluation participants, as more conventional summative evaluation practices often do. The process embodied characteristics of programme coherence that Fullan (2000) suggests is a key component of organizational capacity. One assistant superintendent noted, "The evaluation flew below the radar." That is quite remarkable for an evaluation that continued for more than a year and involved so many people.

Members of the steering committee reported unanimous confidence in the process. The only concern expressed at the final meeting was that they wanted to have been more involved. For a committee of very busy people---three parents, a high school student, three principals, two teachers, and a corporate vice-president---their desire for greater involvement was strong affirmation of the value they attached to the project.

In the early stages, team members were comfortable with the term formative evaluation (Scriven, 1967) to describe the process. As their expertise and confidence grew, they were attracted to Hopkins' (1989) and Fetterman's (2001) concept of empowerment evaluation, and Cousins and Earl's (1995) participatory evaluation. Ultimately, however, team members concluded that they had participated in a form of evaluation that was generative. This term was first adopted during the project to describe a particular type of community, elements of which were identified in the learning communities rubric developed by the action team. In that text, "generative was selected as the word that best captures the spirit of enthusiasm, achievement, mutual respect, and dynamism found in organizations embarked on journeys of learning and discovery" (Townsend & Adams, 2003, p. 15). Later, action team members were able to draw consistent connections between the descriptors of a generative community and the documented effects of the evaluative process on participants and on system operations.

### **Distinguishing Characteristics**

More traditional forms of evaluation give primacy to the role of the external expert. Alternately, the generative evaluation approach is distinguished by its emphasis on:

- multiple ways of knowing
- the learning of all participants
- the value placed on relationships, mutual respect, and trust
- the purposeful linking of the process to established mission statements, principles, goals, and values
- the transparency and accessibility of the process
- the timely and ethical use of new knowledge created through the process
- the pace of implementation
- internal ownership of both the process and the results.

The generative evaluation process that emerged from this project attended to relationships and evidence with equal rigor. Action team members honored the wishes of the school board not to engage in a "brag session" or a "whitewash." As well, they were so diligent in matters of anonymity and confidentiality that not a single complaint about the evaluation was received by team members, school board members, other administrators, or steering committee members. Team members acted confidently on the deputy superintendent's admonition to present all the evidence and, as they honed their skills, they developed a deeper respect for the process and for each other.

Still, the story of this experience would not be complete without a final anecdote. In the weeks leading up to the evaluation, a concern was raised that the cost of the project would take money out of classrooms. Almost immediately, the superintendent responded with a quiet assertion that no such thing would happen. He said, "As president of [a provincial organization] this year, I receive a stipend. I'll ask the school board to use that money to fund the project". The concern was never raised again.

## Conclusion

This early experience with generative evaluation has produced some hopeful results in one jurisdiction. Its success was clearly dependent on many contextual factors such as the willingness of leaders to take risks and share their authority; extant levels of trust; the adequacy of expertise; and the availability of resources. Effective implementation of a generative model may be equally dependent upon the type and focus of the question that guides the investigation, the sincerity of participants, the quality of pre-existing and evolving relationships, and the levels of support for innovation and improvement in the larger educational system.

As a comprehensive model of evaluation, the generative approach complements the principles of learning communities. It has the potential to create and sustain organizational learning while providing increased opportunities for collaboration, recognition, and celebration. It is a model that demands joint responsibility, built as it is on disciplined inquiry, trust, commitment, and mutual respect.

However, while it is a financially responsible model, the generative approach requires much more time. As well, despite its involvement of large numbers of internal team members, it can nevertheless be the cause of conflict if it appears that participants have been unfairly selected. It can promote conflict in other ways, too, particularly if it challenges existing power and authority relationships. It could fail if, in its early stages, data reveal concerns or shortcomings with which the district leadership is not ready or willing to contend. Another risk is that the generative approach could become a self-congratulatory exercise should the integrity of its commitment to ethical uses of all evidence be compromised. Similarly, if the relationship between external and internal team members were to suffer from over-identification (Glesne, 1999), the quality of the evaluation could be eroded.

Further research is required to determine the extent to which the process of generative evaluation has applicability to other contexts. For example, could the generative process be used to determine the respective merits of computer software, the impact of a new teaching strategy, the comparative effectiveness of secondary schools, or the reasons for changes in student achievement? Results of this initial study suggest that the generative process can be an appropriate choice for organizations faced with the need for evaluation, yet wanting to preserve and protect hard-won gains in the quality of relationships among their members.

## References

- Barth, R. (2001). *Learning by heart*. San Francisco: Jossey-Bass.
- Cousins, J., & Earl, L. (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.
- Deal, T. & Peterson, K. (2000). *The leadership paradox: Balancing logic and artistry in schools*. San Francisco: Jossey-Bass.
- Elmore, R. (2000). *Building a new structure for school leadership*. Washington: Shanker Institute.
- Fern, E. (2001). *Advanced focus group research*. London: Sage.
- Fetterman, D. (2001). Empowerment evaluation. *Evaluation Practice*, 15, 1-15.
- Fitzpatrick, J., Sanders, J., & Worthen, B. (2004). *Program evaluation: Alternative approaches and practical guidelines*. Boston: Pearson.
- Fullan, M. (2000). The return of large-scale reform. *Journal of Educational Change*, 2(1).
- Gay, L., & Airasian, P. (2003). *Educational research*. New Jersey: Pearson.
- Glesne, C. (1999). *Becoming qualitative researchers*. New York: Longman.

- Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Thousand Oaks: Sage.
- Hopkins, D. (1989). *Evaluation for school improvement*. Buckingham: Open.
- Lambert, L. (1998). *Building leadership capacity in schools*. Alexandria: ASCD.
- Mertens, D. (1999). Inclusive evaluation: Implications of transformative theory for evaluation. *American Journal of Evaluation*, 20.
- Neuman, L. (1997). *Social research methods: Qualitative and quantitative approaches*. Toronto: Allyn and Bacon.
- Posavac, E., & Carey, R. (1997). *Program evaluation: Methods and case studies*. New Jersey: Prentice-Hall.
- Scriven, M. (1967). The methodology of evaluation. In R. Stake (Ed.), *Curriculum evaluation* (pp. 39-83). Chicago: Rand McNally.
- Senge, P. (1990). *The fifth discipline: The art and practice of a learning organization*. New York: Doubleday.
- Stake, R. (1975). *Evaluating the arts in education: A responsive approach*. Columbus: Merrill.
- Townsend, D., & Adams, P. (2003). *Exploring your learning community*. Innisfail: Chinook's Edge School Division.
- Wengraf, T. (2001). *Qualitative research interviewing*. London: Sage.

Pam Adams  
Faculty of Education  
University of Lethbridge, Alberta, Canada  
[adams@uleth.ca](mailto:adams@uleth.ca)

David Townsend, Ph.D.  
Faculty of Education  
University of Lethbridge, Alberta, Canada  
[david.townsend@uleth.ca](mailto:david.townsend@uleth.ca)