

**DRAFT**

**An Evaluation of the Project:**

*Leading Student Achievement: Our Principal Purpose*

Kenneth Leithwood and Doris Jantzi

OISE/University of Toronto

Paper presented at the annual meeting of the Canadian Society for Studies in Education

Toronto, York University

May, 2006

## **An Evaluation of the Project:**

### **Leading Student Achievement: Our Principal Purpose**

#### **Introduction**

This paper briefly describes the evaluation of a project sponsored by Ontario's Literacy and Numeracy Secretariat. Undertaken in collaboration with the province's three principal associations, the project aims to enhance the literacy and numeracy of elementary school students in the province by further developing the instructional leadership capacities of school principals. Initially funded by the Secretariat for a one year period (2005-2006), the project includes a number of professional development experiences for approximately 60 lead principals, each of whom works with teams of principals in their own districts. Each of these principals, in turn, is expected to work with teams of staff in their own schools helping to create local professional learning communities. Some professional development experiences are common to all participants and entail formally structured sessions on key topics. But the bulk of the professional development is intended to take place within districts and schools through the collective interaction and decision making of those participating principals and their staffs in each district.

The evaluation aims to assess the extent to which the goals of the project are achieved, as well as the relative contribution to such achievement of the methods used in the project. About the achievement of project goals, the evaluation asks:

- How was the leadership of principal participants intended to change?
- What changes actually occurred in the leadership practices of participants?

- What changes actually occurred in the literacy and numeracy instructional practices of participants' teacher-colleagues?
- What changes, if any, occurred in the literacy and numeracy achievement of students in participants' schools?

With respect to the methods used in the project to achieve its goals, the evaluation asks:

- What types of learning opportunities were provided to participants?
- How effective were the methods used to provide these learning experiences?
- How might the effectiveness of these methods and opportunities be improved?
- How much did participants learn from each of these experiences?
- How effective were these experiences in improving principals' leadership practices?

### **Framework for the Evaluation**

The project, as distinct from its evaluation, is based on a professional learning community model with learning teams created at the provincial, district and school levels. Reflecting a broad consensus in the school leadership literature (e.g., Hallinger & Heck, 1996; Leithwood & Jantzi, 2005), evaluation of the project assumes that leadership effects on students' literacy and numeracy achievements are primarily indirect. Viewed as a series of cause and effect relationships, the framework for the evaluation begins with the leadership development experiences provided to participants in the project. These experiences are assumed to influence principals' knowledge, skills and dispositions - especially those considered relevant to the domain of literacy and numeracy.

Of particular interest in this project are the knowledge, skills and perhaps the dispositions associated with improving classroom instruction in literacy and numeracy. Acquisition of such knowledge, skills and dispositions by principals, it is assumed, will

lead to a change in their overt leadership practices or behaviors. These changed leadership practices are expected to stimulate changes in teachers' instructional practices which in turn improve the learning of students. The effects of changed leadership practices are likely to influence teachers' instructional practices both directly and through the creation of professional learning communities.

### **Leadership Development Experiences**

The model for professional development guiding the project combines the use of relatively centralized and formally planned professional development experiences with a range of more informal, local experiences. Within districts, principals work with other administrative colleagues involved in the project to determine and engage in the most useful experiences for themselves. These opportunities reflect evidence about the value of more locally controlled and situationally sensitive professional development and emerging evidence demonstrating the significant impact of professional networks (e.g., Tarleton, 2006).

### **Principal Leadership**

The starting point for what such leadership practices should be is a model of successful "core" leadership practices widely supported by empirical evidence (e.g., Leithwood & Riehl, 2005; Leithwood et al, 2004). The model encompasses four broad categories of leadership practices including *Setting Directions*, *Developing People*, *Redesigning the Organization* and *Managing the Instructional Program*. Within each of these broad categories are some 14 more specific sets of practices. This model of leadership has been modified for the project to reflect the principal associations' and Secretariat's commitment to fostering team, shared or "distributed" leadership (e.g.,

Gronn, 2003; Spillane, 2005). Improved or refined leadership practices, conceived of in this way, are expected to stimulate changes in teachers' classroom instruction.

### **Professional Learning Community (PLC)**

The term "professional learning community" signifies not only discrete acts of teacher sharing, but the establishment of a school-wide culture that makes collaboration expected, inclusive, genuine, ongoing, and focused on critically examining practice to improve student outcomes. School administrators, in particular, help develop PLCs through their attention to individual teacher development, and by creating and sustaining networks of conversation in their schools around issues of teaching and learning. Some recent research using PLCs as a variable has demonstrated powerful associations between PLCs and teacher practices (Bryk, Camburn, & Louis, 1999; Louis, Marks, & Kruse, 1999; Pounder, 1999; Scribner, Cockrell, Cockrell, & Valentine, 1999; Toole, 2001). Reported associations with student learning are rare, however.

As we have noted about PLCs elsewhere (Leithwood, Louis, Anderson & Wahlstrom, 2004):

*The term integrates three robust concepts: a school culture that emphasizes professionalism is "client oriented and knowledge based" (Darling-Hammond, 1990); one that emphasizes learning places a high value on teachers' inquiry and reflection (Toole, 2001) and one that is communitarian emphasizes personal connection (Louis et al., 1995). The hypothesis is that what teachers do together outside of the classroom can be as important as what they do inside in affecting school restructuring, teachers' professional development, and student learning (Louis and Kruse, 1995).*

Kruse, Louis, & Bryk (1995) identified five dimensions of a genuine professional learning community: shared norms and values, a focus on student learning, de-privatized practice, reflective dialogue, and collaboration. Our evaluation focused primarily on the de-privatized practice dimension since it is associated with the strongest effects on changed teacher practice.

### **Instructional Practices**

Determining the most productive forms of instructional practice in language and math has been a large part of the Secretariat's work to date; it is the ongoing focus of those members of the Secretariat responsible for teacher professional development in the province. Practices encouraged by the Secretariat are described in the work of the Expert Panel<sup>1</sup>, the Ontario curriculum and guides produced by the Ministry of Education and a range of research evidence about effective literacy and numeracy instruction. Viewed broadly, these practices reflect what is more loosely described as "constructivist" forms of teaching (Marlowe & Page, 2005) - teaching aimed at building understanding in both language and math (e.g., Barr, 2001; Ball et al, 2001). The *Literacy for Learning* report (2004), for example, encourages teachers to integrate language instruction across the curriculum, to "...use questioning and inquiry to engage students in higher order thinking and critical literacy practices as they read and write" (p. 27) and "...to create 'conversational communities' where talk is central to learning" (p. 28).

### **Student Learning**

Changes in teachers' instructional practices are expected to result in improved student literacy and mathematics learning. These improvements in learning are being estimated through the use of provincial test results from students in grades 3 and 6.

---

<sup>1</sup> Report of the Expert Panel (2004). *Literacy for Learning*.

## **Evaluation Methods**

### **Design**

A modified pre-test, post-test design is being used for the project. The modification entails creating comparison groups in order to estimate the effects of the project on student achievement. Among the likely comparisons to be made are: changes in achievement in project participants' schools vs. changes in achievement in all other schools in the district; changes in achievement in participants' schools vs. other schools in the district matched on a series of variables including student SES and aggregated mean achievement at the time of project initiation.

### **Instruments and Sources of Data**

Data for the project are being collected from teachers, principals and supervisory officers and students. Some documents are also to be collected and the content analyzed.

*Teachers.* Data is being collected from teachers using an on-line survey instrument created through significant modifications and additions to already well-developed, reliable instruments. The survey includes multi-item scales measuring school leadership, general approaches to instruction, specific instructional practices used to teach both language and math and professional learning community. The survey also includes questions asking about a small number of demographic variables.

Scales measuring instructional practices in language and math reflect the practices described in sources mentioned above; they were also the subject of considerable review and modification by Secretariat staff. The survey was administered first in October 2005 and will be administered a second time in May 2006.

*Principals and other administrator participants.* Data from principals and superintendents involved in the project has been collected through a series of phone interviews conducted during January and February 2006. These interviews have inquired about the nature and perceived value of the professional development experiences in which participants have engaged both locally and provincially. This interview data will be supplemented with survey data toward the end of the project.

*Students.* Achievement data in math and language are being provided through tests administered by the Educational Quality and Accountability Office to students in grades 3 and 6. Two estimates of achievement have been calculated. One estimate is an average annual achievement score for each school in each of grades 3 and 6 (the proportions of students scoring at or above the proficiency cut score on both language and math exams). The second estimate is a growth measure - not only growth over the one year duration of the project (spring 2005 to spring 2006) but over the prior four years, as well.

There are, of course, serious flaws in the evaluation design if impacts on students are to be assessed by estimates of growth only over the current one-year project funding period. For example, it is ambitious, in the extreme, to expect detectable changes in student achievement over less than one year of project activity. Furthermore, changes in student scores across a school over one year are very unstable and the causes very hard to determine (e.g., Linn, 2003). These are widely recognized flaws that will be minimized should there be opportunities to continue the project beyond its current one year period.

## **Sample**

At the time of preparing this paper, we were unable to determine the size of the intended sample, although this information should be available shortly. The achieved sample included a total of 522 teachers (205 public, 208 Catholic and 109 French) in 110 schools (53 public, 57 Catholic and French). For the 70 schools with complete data sets, 25 were public (118 teachers) and 45 were Catholic (173 teachers).

## **Some Preliminary Results**

### **Three Sets of Relationships**

Although the primary uses to be made of the teacher survey and student achievement data are to estimate change over the one year period for which the project is currently funded, the pre-test survey data along in combination with EQAO achievement data speak to several questions important in their own right, questions not anticipated in the original design of the study. Specifically, an analysis was carried out to explore the relationships between student achievement and (a) teachers' instructional practices, (b) team leadership practices and (c) professional learning communities.

### **The Preliminary Analysis**

Survey results from individual teachers were used to calculate values for five variables measured with the on-line survey: literacy instruction, mathematics instruction, general approaches to instruction, team leadership practices and professional learning community. The value for each variable was obtained by calculating the mean for the items making up the scale used to measure the variable. Teacher data were then aggregated to the school level for further analyses. At this point, none of the data from the French language schools have been aggregated to the school level so the analyses are

limited to the schools of principals represented by the *Ontario Principals' Council* and the *Catholic Principals' Council of Ontario*.

Student achievement data for the schools with survey responses were obtained from the EQAO website for grades 3 and 6 for the 2001-02, 2002-03, 2003-04, and 2004-05 academic years. The percentage of students at or above the provincial standard in reading, writing and mathematics in each of these years was added to the aggregated data file. A single language score was obtained by calculating a mean from the reading and writing scores.

A total of 89 schools had a complete set of grade 3 language scores and 90 had grade 3 mathematics scores. A complete set of grade 6 language scores were available for 81 schools and grade 6 mathematics scores for 81 schools. Sixteen sets of achievement data were available: achievement scores in two subjects (language and mathematics) for four years. In addition, four gain scores were calculated (grades 3 and 6 math and language) by summing gains in achievement each year from 2002 to 2005. Correlation coefficients were calculated for a series of relationships between each of our estimates of student achievement and survey variables in the 70 schools with data for all achievement and survey measures.

### **Achievement and Instructional Practices**

Our analysis indicated no significant relationships between literacy and mathematics instruction and gains in student achievement over four years. Literacy instruction had a weak ( $r = .25$ ) relationship with grade 3 language achievement in 2003 and mathematics instruction had a weak relationship with mathematics achievement in the same year ( $r =$

.27). General instruction was related to grade 3 language and mathematics achievement in 2003 (both  $r = .30$ ).

More refined analyses were done by correlating individual literacy and math strategies with the eight annual achievement scores for that discipline. Over half the literacy strategies (9) were not related to any achievement data. Two strategies were related positively to one score and one strategy to two scores. Three strategies (“answer questions that require inferences”, “engage students in peer evaluation”, and “reflect on their own learning strategies”) were each significantly, but weakly, related to three achievement scores. The strategy, “Identify the author’s purpose” was related to four scores. Of the 17 mathematics strategies, 13 strategies were not related to achievement and four were related to one achievement score. None of the instructional strategies were related to gains in achievement from 2002 to 2005.

### **Achievement and Team Leadership**

The most significant relationships (seven) were between team leadership practices and achievement. Leadership was the only variable related to achievement in grade 6. Leadership was related to grade 3 language in 2003 ( $r = .42$ ), grade 3 mathematics in 2003 ( $r = .30$ ), grade 6 language in 2003 ( $r = .33$ ), 2004 ( $r = .34$ ) and 2005 ( $r = .33$ ), and grade 6 mathematics in 2003 ( $r = .23$ ) and 2004 ( $r = .27$ ).

The 20 team leadership practices were correlated with all 16 achievement scores and four gain scores. Two practices were not related to achievement, 11 practices were related to four or fewer scores; and three practices were related to five scores. “There is a sense of overall purpose in this school” was positively related to six scores, but somewhat puzzling, because it was weakly and negatively related to grade 3 language achievement

gain. Two practices (“There are high expectations for staff activities with students” and “An atmosphere of trust and caring exists”) were significantly related to half the achievement measures. “Parents are engaged in school improvement efforts” was significantly related to 11 of the achievement scores. Finally, “New ideas and strategies for professional practice are encouraged” was significantly, albeit weakly, related to gains in grade 6 language achievement ( $r = .24$ ).

### **Achievement and Professional Learning Community (PLC)**

PLC was not significantly related to any of the achievement measures.

### **Conclusion**

Our pre-test measures of literacy and mathematics instruction appeared to be largely unrelated to student achievement calculated either as an annual achievement score or as a growth over four years. These are instructional practices consistent with the province’s Expert Panel, it provides curriculum and local interpretations of research about best practices. These practices are the focus of much of the teacher professional development being provided throughout the province. Furthermore, the measures were carefully reviewed and approved by members of the Secretariat closely involved in these areas of instruction.

In contrast, our measure of effective team leadership appears to be consistently and significantly related to student achievement. Most frequently, this relationship is with achievement calculated as an annual achievement score, but in some cases also as a change-over-four-years score. In a “horse race” between classroom instruction and leadership to determine which has the greatest impact on achievement, classroom

instruction normally would be expected to win, hands down. So this set of relationships is quite surprising.

As usual, this data is not perfect and some caution needs to be exercised in deciding just how to make sense of the surprises. Most notably:

- Response rates within many schools were quite low; about 60% of schools had fewer than four teacher respondents and only 10% had eight or more. This is a reason to be cautious about the representation of the data at the school level. On the other hand, it would be plausible to imagine that returns would be most likely from the keenest of the language and math teachers;
- There was no way to link teachers to grade levels, so it is not possible to determine whether the respondents had taught students who were tested. This is a reason to worry some about how well the survey data represent what is actually going on in each school. However, teachers at all grade levels contribute to scores for grade 3 and 6 students. So this does not seem a fatal limitation
- Teachers' ratings of their own instructional practices were quite high and standard deviations were quite low. Especially the restricted range of standard deviations reduced the likelihood of finding significant relationships. Nonetheless, the scores for leadership items were also fairly high and standard deviations only slightly larger than was the case for instruction items.

Because the data set for the project are still quite incomplete, it would be premature to draw any firm conclusions. It seems safe to say, however, that the results may hold some surprises, as the preliminary analysis of pre-test data illustrates. These surprises may well speak to the efforts of policy makers in the province now to find the

right balance between allocating resources to further leadership development on the one hand and further instructional development, on the other.

## References

- Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433-456). Washington, DC: American Educational Research Association.
- Barr, M. (2001). Research on the teaching of reading. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 390-415). Washington, DC: American Educational Research Association.
- Gronn, P. (2002). Distributed leadership. In K. Leithwood & P. Hallinger (Eds.), *Second international handbook of educational leadership and administration* (pp. 653-696). Dordrecht: Kluwer Academic Publishers.
- Hallinger, P., & Heck, R. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, 32(1), 5-44.
- Leithwood, K., & Jantzi, D. (2005). *A review of transformational school literature research 1996-2005*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, QC.
- Leithwood, K., & Riehl, C. (2005). What we know about successful school leadership. In W. Firestone & C. Riehl (Eds.), *A new agenda: Directions for research on educational leadership* (pp. 22-47). New York, NY: Teachers College Press.
- Leithwood, K., Seashore-Louis, K., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning: A review of research for the Learning from Leadership Project*. New York: The Wallace Foundation.
- Linn, R. (2003). Accountability: responsibility and reasonable expectations. *Educational Researcher*, 32(7), 3-13.
- Marlowe, B. A., & Page, M. L. (2005). *Creating and sustaining the constructivist classroom* (2nd ed.). Thousand Oaks, CA: Corwin.
- Spillane, J. P. (2006). *Distributed leadership*. San Francisco, CA: Jossey-Bass.
- Tarleton, R. (2006). *Leadership networks add value*, Nexus (Special Edition), Spring (a publication of the National College for School Leadership).

