

An Investigation of the Relationship between Professional Learning Community Practices and Student Achievement in an Eastern Canadian School Board

Noel Hurley, Tim Seifert, & Bruce Sheppard
Faculty of Education, Memorial University

Abstract

The participants in this study were teachers and students of a large school district located on the east coast of Canada. We distributed surveys to 1,514 teachers across 84 schools and received 1,423 usable surveys for a return rate of 94%. At the time of the second data collection, three years later, 1574 teachers from 78 schools agreed to participate. Teachers' perceptions of their schools as professional learning communities were assessed using 32 items in a 5-point Likert format. The responses to the 2006 survey were factor analyzed using ML estimation. A number of different factor structures were tested, including a single factor model and a second-order factor analysis. In spite of a heavy emphasis by school board leaders and the collaborative development, introduction, and implementation of PLC practices before and during the study period, no significant measurable effect on student learning was identified.

Keywords: professional learning communities, change implementation, educational change, collaborative leadership, linear growth models, and reading achievement

In the past decade or so, a growing emphasis on student achievement as a measure of school success has contributed to an increased research focus on the determination of the effects of school leadership upon student test scores. At this point in time, most researchers have come to recognize that developing any meaningful direct connection between the role of the formal leader (e.g., the school principal) and student learning outcomes is unlikely (Anderson, Moore, & Sun, 2009; Leithwood & Mascal, 2008; Mascal, Leithwood, Strauss, & Sacks, 2009). As a consequence, most researchers have focused on identifying both the direct and indirect paths by which various leadership sources and/or various leadership functions influence important school goals (York-Barr & Duke, 2004). This shift in research focus in respect to school leadership has resulted in growing evidence that leadership is more effective when it is shared among professionals within the school (DuFour, Eaker, & DuFour, 2004; Elmore, 2000; Marzano, Waters, & McNulty, 2005; Reeves, 2008). Menlo (2011) contended that not only is shared decision making important in bringing about beneficial improvements to schooling, but it is essential that those with whom decision making is shared take up the responsibility to take action for the sustainability of those changes. Similarly, Hargreaves and Fink (2003) stated that wise principals realize that they cannot reach instructional goals on their own. They contended that principals who distribute leadership over their schools are more likely to sustain school improvement. Leithwood, Seashore-Louis, Anderson, and Wahlstrom (2004) claimed that the most successful principals develop and rely on the expertise of teacher-leaders within their schools.

Consistent with the above noted evidence related to the importance of distribution of leadership for school success, over the past decade many school jurisdictions, particularly in the western world, have in-

vested much time and money in adapting what have become known as Professional Learning Community (PLC) practices (DuFour et al., 2005; Giles & Hargreaves, 2006; Hall & Hord, 2006). Even though many school jurisdictions have adopted these PLC practices as a means of improving student learning, research evidence revealing the success of this approach to improvement remains sparse. The paucity of empirical evidence revealing a positive impact of professional learning community initiatives led Sheppard, Brown, and Dibbon (2009) to observe that “the term *professional learning* has been co-opted by schools, school districts, departments of education, and consulting firms as the current label of choice (the buzzword of the day)...” (p. 11) that offers little more than a simplistic recipe for immediate improvement that is bound to disappoint. It appears that others (e.g., Hall & Hord, 2006, and Dufour, 2004) shared similar concerns. For instance, Hall and Hord (2006) pondered whether it has become the “innovation du jour” (p. 270) whereby many schools announce that they are PLCs without having any of the attributes or characteristics that have been identified as most valued in the empirical literature. Similarly, Dufour (2004) has lamented that “the term has been used so ubiquitously that it is in danger of losing all meaning” (p. 6).

From our perspective, a school cannot be a PLC unless there is a school culture that fosters teacher leadership (Crowther, Kaagan, Ferguson, & Hann, 2009; Sheppard & Dibbon, 2011). The ability of a school to learn “is dependent on [its] capacity...to facilitate collaboration among individual learners [informal leaders such as teachers] who assume distributed leadership responsibilities and learn from one another” (Sheppard et al., 2009, p. 16). It is, therefore, the role of the formal leaders to facilitate teacher leadership by being transformational and inclusive, and by providing opportunities for professional learning and collaborative engagement.

In response to the apparent gap between the empirical evidence and actual practice in schools and school districts, we undertook a district-wide study in one school district in Atlantic Canada to investigate the relationship between PLC practices and student learning. We sought to determine if a focused district-wide initiative to increase the use of PLC practices would result in improved student learning as has been so widely accepted by many policy makers.

Method

Participation

The participants in this study were teachers and students of a large school district located on the east coast of Canada. As a result of a partnership arrangement with the school district, our survey return rates for both administration periods were excellent. For the 2006 data collection, we distributed surveys to 1,514 teachers across 84 schools and received 1,423 usable surveys for a return rate of 94%. At the time of the second data collection, three years later, 1,574 teachers from 78 schools agreed to participate. The return rate during the second data collection was 81%. It is important to note that some very small schools did not participate in the first data collection because of concerns with anonymity. However, they chose to participate in the second data collection.

The school district was comprised of a mix of rural and urban, and small and large schools. As evident in Table 1, there was an increase in the number of schools with small numbers of Grade six students completing the assessment and a decrease in the number of schools with large numbers of students (indicative of a decline in enrollments).

Table 1. Descriptive statistics of standardized Provincial Literacy Assessment and school sizes in the school district.

	2006	2007	2008	2009
Mean	509.01	496.41	496.65	501.50
<i>SD</i>	94.09	102.84	103.59	102.90
Minimum	200.00	200.00	200.00	200.00
Maximum	603.00	639.00	639.00	647.00
<i>N</i>	2,882	1,701	1,649	1,537
% Meeting criteria	.91	.83	.87	.87
Number of schools with:				
<10 Gr. 6 students	1	7	8	9
11-20 Gr. 6 students	5	7	9	5
21-30 Gr. 6 students	9	11	9	10
31-40 Gr. 6 students	10	4	4	5
41-50 Gr. 6 students	9	5	4	3
>50 Gr. 6 students	24	9	9	10

Procedures

The Professional Learning Community survey. Data were collected through the administration of a survey during two collection periods separated by a period of three years. Our survey instrument, adapted from previous work (Sheppard & Brown, 2000), was designed to assess collaborative school leadership and professional learning community characteristics. The questions included in this survey are based on an extensive review of the theoretical literature related to organizational learning, professional learning communities, and collaborative leadership. We have used it extensively for more than a decade (Sheppard et al., 2009) and have made revisions to individual items and constructs on the basis of new empirical evidence and factor analyses of data from multiple collection events. As well, additional evidence in support of the content and construct validity of this instrument has been gathered in more than two-dozen organizational/professional development sessions with school system leaders, school principals, and teachers with whom we have established research and development partnerships. During these aforementioned sessions, each of the survey items and constructs have been subjected to considerable scrutiny by our partners in the school system (our survey respondents) and modified as appropriate.

Data were analyzed using MPlus v7 (Muthén & Muthén, 1998-2012) and SAS v9.3 (SAS Institute, 2010) software. Teachers' perceptions of their schools as professional learning communities were assessed using 32 items in a 5-point Likert format. The responses to the 2006 survey were factor analyzed using ML estimation. A number of different factor structures were tested, including a single factor model and a second-order factor analysis. The final solution, based upon the c^2 , Tucker-Lewis Index, Bayesian Information Criteria and Root Mean Square Error of Approximation fit indices (Garson, 2009; Hu & Bentler, 2000), was a five-factor solution (see Table 2 and Figure 1). These PLC factors represent:

- The degree to which the school is perceived as democratic (Democratic environment);
- The degree to which the school is perceived as collaborative (Collaborative culture);
- The degree to which teachers in the school collaborate (Teacher collaboration);
- Teachers' sense of being able to influence student learning (Teacher efficacy and confidence); and
- The extent to which teachers feel valued (Feeling valued).

Table 2. Results for a set of factor analyses models of teacher responses the PLC survey items.

	χ^2 (df)	CFI	TLI	RMSEA	SRMR
Single unitary factor	4587.21 (350)	.74	.72	.092 (.090, .094)	.076
Five 1 st order factors	1660.11 (340)	.920	.911	.052 (.050, .054)	.049
Five 1 st order factors, 1 2 nd order factor	1772.03 (345)	.913	.911	.052 (.050, .054)	.049

Missing values for factor scores at the teacher-level were assigned the mean value for the respective factor. The resulting factor scores from teacher-level data were then aggregated to the school level by computing a PLC score for each school. The factor scores were then rescaled to be centered on the mean of the items constituting the factor, with a spread equivalent to the mean variance of the items. By rescaling the factor scores, the means and standard deviations are readily interpreted relative to the Likert scales from which they were derived. This factor structure was then imposed upon the 2009 teacher responses. As with the 2006 data, missing values for factor scores were assigned the mean value for the respective factor, and the factor scores rescaled to be centered upon the mean of the corresponding items.

The internal consistency coefficients of five of the factors were good (see Figure 1). The factor loadings were moderate to strong. However, strong correlations among several of the factors were noted. The democratic environment and collaborative culture factors were strongly correlated ($r=.88$). The democratic environment, collaborative culture, and teacher collaboration factors were strongly correlated with the feeling valued factor. While the democratic environment and collaborative culture factors were strongly correlated with the collaborative school leadership factor, the other factors were weakly or moderately correlated with it (see Figure 1).

Student reading achievement. Student achievement in reading was measured by a series of criterion-referenced tests as part of the Province's assessment of literacy and mathematics. The tests, the Elementary Literacy Assessment, were developed by the Province's Department of Education and were administered to all Grade 6 students in the Province (Atlantic Canada English Language Arts Curriculum Guide: Grades 4-6). The assessments were administered on the same four days during the first hour and a half of school at the same time in every school throughout the Province. These assessments were administered in October and scored in November with individual reports being sent to schools and then distributed to parents and guardians by mid-winter. To ensure an accurate picture of students' independent literacy skills, teachers did not help students other than to direct them to the appropriate pages on which to work. Student seating was arranged to ensure that students had privacy when participating in the assessment.

Teachers and consultants, who were familiar with the curriculum outcomes and trained to assess them, from all provincial school boards were employed to score them. The assessment was designed to assess specific curriculum outcomes that students were expected to have mastered by the end of Grade 5. The assessment addressed only those outcomes considered suitable for an external assessment; namely, those outcomes associated with reading, viewing, and writing domains (Atlantic Canada English Language Arts Curriculum Guide: Grades 4-6). Thus the reading results that we used were a reliable representation of student achievement in reading among students entering Grade 6.

Although both reading and writing components were assessed, only the reading achievement data were analyzed for this study. For this portion of the test, students were asked to read and answer questions on materials that included expository texts, short narratives, poetry, and visual media. The questions assessed comprehension and analysis, and curricular outcomes such as use of reading and linguistic cuing strategies, recognition of conventions, and thinking critically about the text.

Test scores were scaled to a provincial average of 500 with a standard deviation of 100. Because they were standardized on a yearly basis, interpretation of the scores must be cautious. The scores are normative; over the course of the four-year period the average score in the Province will show no growth because the scores were centered on 500 each year. Scores must be interpreted relative to the provincial average and trends. An increase (or decrease) in a school's achievement scores means that that school improved (or declined) relative to everybody else and the provincial average.

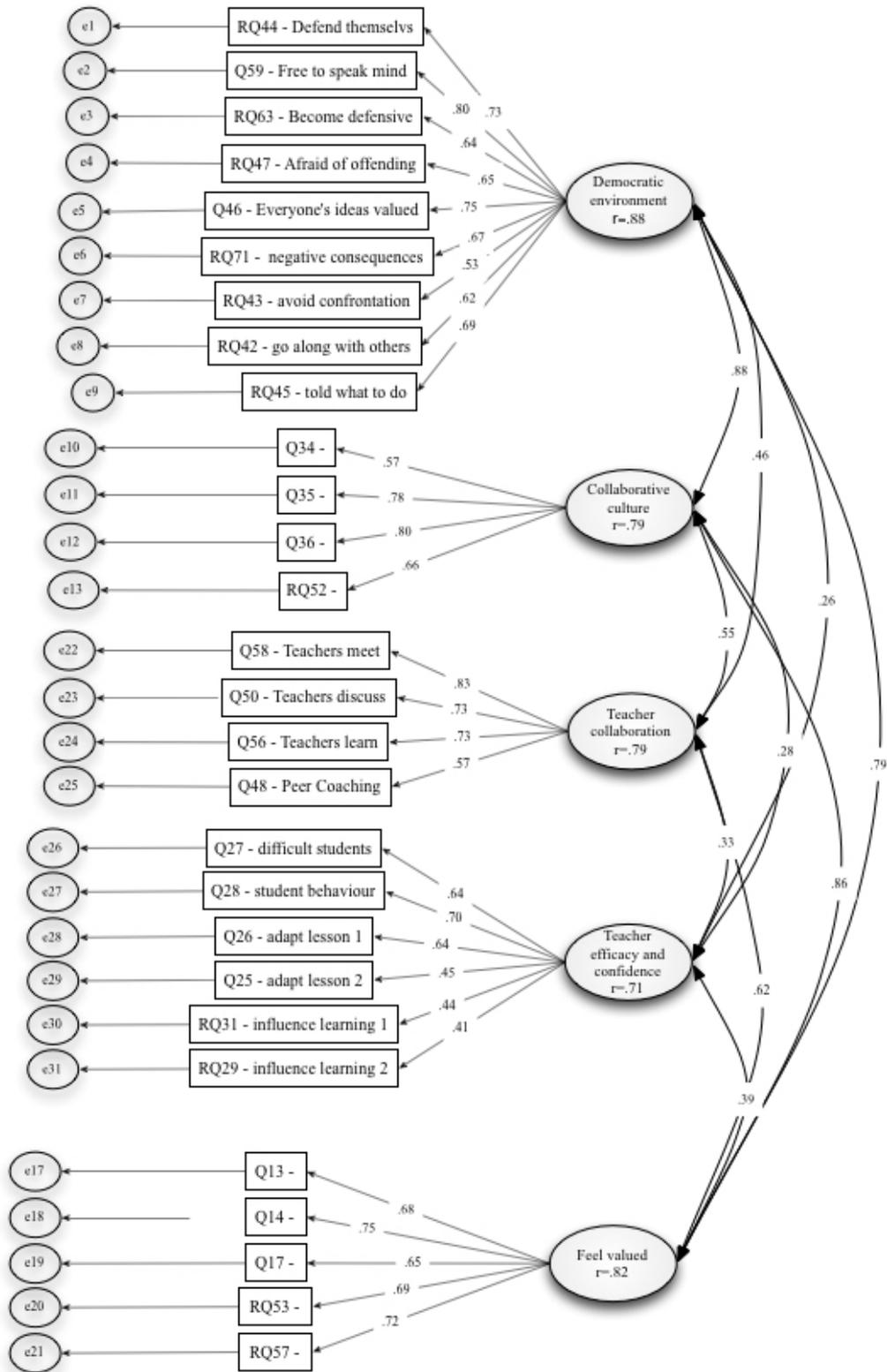


Figure 1. Factor structure of Professional Learning Community items.

Implementation of a Professional Learning Community

In 2006, the District undertook a concerted effort to develop professional learning communities in its schools. These efforts were multi-faceted and were implemented throughout the study period. The approaches included limiting the number of goals that were being addressed by the Board's strategic plan and by explicitly stating the increase of the use of PLC practices as one of the Board's three main goals. In keeping with PLC practices, the Board also placed a greater emphasis on increasing student learning, specifically in the areas of literacy and numeracy. During and after 2006, the Board provided district-wide professional learning opportunities to all personnel. Many school- and board-level leaders attended national and international conferences that were designed to increase participant knowledge of PLCs and that provided training to enable these leaders to train others within the system. Several internationally known experts on the implementation of PLCs were also brought to the school district to provide direct learning opportunities to teachers, consultants, directors, principals, and board members. To pursue the goal of continuous improvement, many board employees participated in professional development sessions with leaders associated with professional learning research such as Ken Leithwood, Rick Dufour, Rebecca Dufour, Bruce Sheppard, Robert Eaker, Wayne Hulley, Andy Hargreaves, Dean Fink, Dan Reeves, Avis Glaze, Richard Marzano, and Michael Fullan.

Results

The initial plan for data analyses called for a four-step process. The first two steps examined the growth (or decline) of levels of PLC practices over a three-year period. The first step examined changes in teachers' perception of their school as a professional learning community. In the second step we examined schools as professional learning communities. Third, we modeled growth in reading achievement over the three-year period, and finally, we linked change in PLC practices to reading achievement.

The analyses for the first two steps proceeded as expected and we were successful in modeling the change in PLC scores over the three-year period. However, unexpected problems were encountered in modeling the achievement levels in schools over time. Attempts to model achievement growth in schools were met with the problem of non-positive definite matrices. While this problem may be an artifact of the small sample size (25 schools with students who participated in the testing in all four years as there were a limited number of schools who offered programs at the elementary grade level), it is more likely the result of the inconsistent patterns of achievement displayed within schools over the three testing periods may have led to instability in the variance-covariance matrices. When convergence was achieved, the resulting model fit the data poorly, and the low number of schools in the sample led to low power.

Consequently, the plan for the analyses changed. Following the modeling of teacher and school PLC scores, we proceeded with a series of case studies in which we compared and contrasted the achievement levels of schools with similar patterns of PLC growth. Three questions guided the case selection and analyses:

1. If a school's PLC score increased, what was the resultant effect on achievement?
2. If a school's PLC score decreased, what was the resultant effect on achievement?
3. If a school had a consistent PLC score, what was the resultant effect on achievement?

Professional Learning Communities in the Schools: Teachers' Perceptions

From the data, we were able to match the responses to the 2006 survey with responses to the 2009 survey for 317 teachers in 51 schools. Of those 51 schools, approximately 63% had fewer than five teachers while approximately 17% had 10 or more teachers. The results of attempts to match 2006 surveys with 2009 surveys resulted in four groups. The first group contained no identification codes, but PLC scores for 2006 and 2009. The second group contained identification codes and PLC scores for 2006 but not for 2009 (missing data). The third group had identification codes and PLC scores for 2009 but not for 2006 (missing data). The last group had identification codes and PLC scores for both 2006 and 2009. An omnibus between-groups test of 2006 PLC scores yielded a statistically detectable effect, and Tukey's post-hoc tests indicate differences between specific groups, calculated effect sizes of mean differences were small (maximum $d=.09$). This suggests that the group of teachers for whom we could match 2006 and 2009 PLC scores did not differ on PLC scores from those who had missing data. Descriptive statistics for the four groups, as well as the results of the ANOVA are presented in Table 3a.

Table 3a. Descriptive statistics for missing data, useable data, ANOVA of missing data.

	PLC 2006		PLC 2009		<i>N</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>	
No TID, PCL 2006 and PCL 2009	18.20	.94	17.72	1.95	686
TID, no PLC 2006, PCL 2009	--	--	17.97	1.98	275
TID, PLC 2006, no PLC 2009	18.14	2.06	--	--	925
TID, PLC 2006, PLC 2009	18.46	1.80	18.29	2.02	317
	<i>F</i>	4.09	9.10		
	<i>df</i>	2, 1925	2, 1275		
	<i>MS_w</i>	11.78	35.56		
	<i>p</i>	.017	<.0001		

Note: TID refers to a useable identification number that allowed matching to occur. No PLC means that a PLC score could not be attributed to a particular teacher in that year.

The model was expanded to a three-level model in which time was nested within teachers who were nested within schools. The results (see Table 3b) indicated that while there was no overall change and teachers exhibited individual differences, there were school effects indicating that teachers across schools differed in their perceptions from 2006 to 2009. That is, there were differential effects across schools.

Professional Learning Communities in the Schools: School-Level Analyses

The results of the three-level model of the teachers' responses indicated differential changes across schools. However, that model only drew on a subset of data for which 2006 and 2009 could be matched. To expand the analyses and draw on a larger response pool, we aggregated teacher responses to the school level.

Table 3b. Results of multi-level modeling of PLC scores for teachers and teachers within schools.

		Intercept				Slope				
		Mean	<i>p</i>	Variance	<i>p</i>	Mean	<i>p</i>	Variance	<i>p</i>	ρ
Teachers	Time effect	18.46	<.0001	1.95	<.0001	-.17	.105	.81	.01	.60
Teachers within schools	Time effect	18.61	<.0001			-.13	.38			
	School effect			.68	.002			.48	.009	.26
	Teacher effect			1.24	<.0001			.22	.22	.48

The analyses began with two multi-level unconditional means models (e.g., Raudenbush & Bryk, 2002; Singer, 1998). In these models, teacher PLC scores were level one units nested within schools; one model was for 2006 data and the other for 2009 data. Results, presented in Table 4a and Table 4b, show several points. First, the intercepts changed little from 2006 to 2009 (which will be described further on). In both 2006 and 2009 there was a statistically detectable effect for variance, indicating that schools differed in the perceived level of PLC by teachers within the schools. The moderate values for the intra-class coefficient suggest that schools had effects on teachers' PLC scores.

Teacher PLC scores were aggregated within schools by computing the average teacher PLC score. These school-level mean values were then analyzed with a linear growth model (e.g., Raudenbush & Bryk, 2002; Singer, 1998). The resulting model is presented in Table 4b. On average, there was a slight

decline in PLC scores from 2006 to 2009. Noteworthy, however, is that the variance of the slopes was not statistically detectable. That is, while there was a slight decline in PLC scores over time, schools tended to exhibit the same rate of decline. Although the variance in slopes was not statistically detectable ($\tau=.099$, $p=.282$), it is clear that there are schools which showed considerable growth or decline. The lack of a statistically detectable non-zero variance may be an artifact of low power (due to a small number of schools; $n=63$) and the fact that most schools exhibited little change in PLC scores over time.

Table 4a. Results of multi-level unconditional means models of teacher PLC scores.

	Intercept				
	Mean	<i>p</i>	Variance	<i>p</i>	ρ
2006 Teacher PLC scores	18.63	<.0001	1.15	<.0001	.297
2009 Teacher PLC scores	18.16	<.0001	1.37	<.0001	.329

Case study set 1 – growth in PLC and achievement. The next phase of the analyses consisted of a series of cases studies. The first school chosen, School A (see Table 5), could be described as a medium-sized school. In 2006, the school had 16 teachers who completed the PLC survey and 47 students in Grade 6 who completed the Provincial Assessment in 2007. In 2006, there were five students in the school who had received suspensions for offenses ranging from defiance and threats to sexual misconduct and assault. By 2009, the number of teachers who completed the PLC survey had increased to 20, and the number of students completing the survey had grown to 66. Suspension data were not available.

Table 4b. Results of linear growth model of school-level PLC scores.

	Intercept				Slope				
	Mean	<i>p</i>	Variance	<i>p</i>	Mean	<i>p</i>	Variance	<i>p</i>	ρ
Time effect	18.70	<.0001	.76	<.0001	-.54	.001	.25	.23	.53

School A was located in a rural area but was on the edge of a town with a population of about 10,000. Most of the students attending the school came from families that were below the provincial average in income and employment levels. The principal of School A was a young, energetic woman who was enthusiastic and who led using a very student-centered approach to school leadership. The school had a welcoming atmosphere and expectations for student achievement were high. The school building was only two years old and was equipped with state of the art technology and learning spaces. As a result of perceived success in her school, her leadership team was one of only a few schools that were selected from the whole province to speak about collaborative leadership and other learning community practices at the provincial principals' conference in 2009. They demonstrated how provincial assessments can be used to intervene to improve student achievement of assessed outcomes.

In 2006, the mean PLC score of teachers in School A was not different from other teachers in the district ($t=1.07$, $p=.28$). However, the variance of those scores was less than those of the other teachers, indicating that the teachers had more cohesive scores than other teachers ($F=2.43$, $p=.05$). Over the three-year period, School A reported growth in their school as a professional learning community and the mean PLC score exceeded that of the rest of the district ($t=4.71$, $p<.0001$). The variance of the scores increased slightly and was no longer statistically different from the rest of the district ($p=.08$). However, the students in the school demonstrated consistent improvement in literacy. Over the period from 2007 to 2009, the achievement scores improved by .4 standard deviations and exceeded the provincial average ($t=2.79$, $p<.005$), a noteworthy accomplishment because scores are gauged relative to the rest of the province. The number of students who met the minimum criteria was higher than the provincial level, and approached statistical significance ($\chi^2=2.79$, $p=.09$).

School B in the first set of case studies is a small school, with only six teachers completing the PLC survey in 2006. By 2009, this number had dropped to four. The school experienced an increase in the number of students, increasing from 12 students completing the achievement test in 2007 to 20 in 2009. In 2006, the school dispensed eight suspensions for disruption, disrespectful behaviour, and aggression; two students accounted for seven of the eight suspensions. The school was situated in a large town and had a

student body drawn from a relatively homogeneous and stable population. The principal had a high level of respect among staff members and was an experienced leader who modeled effective teaching and learning strategies with student learning as the school's main focus. The school building was more than 100 years old but was very well maintained and with extra space that permitted flexible learning approaches. The leadership team maintained an emphasis on student-centered learning in an atmosphere that sought to continuously improve. In 2006, the PLC score for this school was the second lowest in the district, and was statistically lower than the rest of the district ($t=2.25, p=.02$). By 2009, the PLC scores for this school had risen to close to the district average and were not statistically different than the scores for the rest of the district ($t=.30, p=.76$). Over the three-year period, student achievement scores exhibited growth. The proportion of students meeting the minimum criteria level was .67 in 2007. This was descriptively lower than the student district-wide level of .85. By 2009, the number of students reaching minimum criteria had risen to .86 indicating that the level of literacy in the school was on par with the rest of the district. The 2007 reading scores in School B were .4 of a standard deviation below the district level. This was not a statistically detectable difference, but the lack of a statistically detectable difference is an artifact of low power. Over the three-year period, the students' literacy scores rose to be on par with students across the district.

The final school for the first set of case studies, School C, is similar to School B in some respects. It is a small school with 9 teachers completing the PLC survey in both 2006 and 2009. Twenty-seven students completed the Provincial Assessment in 2007, and by 2009 that number had dropped to 20. In 2006, two suspensions had been handed out for relatively minor offences. The school was located in a small town in a rural part of the province that has seen steadily declining enrollments over the past two decades. There is a strong community emphasis on learning in this area that goes back many decades but parents expect high levels of achievement in the arts as much as they do in core subject areas. The emphasis on areas not assessed in the literacy assessment might possibly have had a depressing effect on the scores received on the provincial assessment.

The average of the 2006 PLC scores in School C was on par with the district level ($t=.85, p=.39$). However, there was considerably more variance in the scores for this school ($\sigma^2=2.75$ for School C versus 1.96 for the district) which was approaching a statistical difference ($p=.09$). By 2009, the average PLC score in School C had risen to the point where it exceeded the district average ($t=3.51, p=.0005$). Also noteworthy, though, is that the variance in the scores had decreased ($F=6.61, p<.01$). This means that over the three-year period, not only did the teachers' perceptions of their school as a professional learning community increase, they became more cohesive in their opinions.

Table 5. Descriptive statistics for the eight schools selected for case study analysis.

		PLC score		Provincial assessment score			
		2006	2009	2007	2008	2009	
Increase in PLC scores	School A	Mean	18.75	20.29	495.49	499.72	535.97
		SD	1.26	1.41	95.33	92.14	86.56
		N	16	20	47	79	66
		Proportion	--	--	.85	.86	.88
	School B	Mean	16.42	17.92	468.03	526.74	500.65
		SD	1.77	1.80	138.37	78.82	83.71
		N	6	4	12	27	20
		Proportion	--	--	.67	.89	.86
	School C	Mean	18.78	20.53	522.81	478.65	502.56
		SD	2.76	1.08	108.95	96.67	122.03
		N	9	9	27	21	20
		Proportion	--	--	.85	.67	.75
Decrease in PLC scores	School D	Mean	19.29	17.21	461.36	454.64	441.96
		SD	.61	.38	93.05	76.76	123.20
		N	5	6	14	11	21
		Proportion	--	--	.79	.75	.60
	School E	Mean	19.16	17.70	527.88	456.43	447.77
		SD	1.34	1.87	82.08	97.03	119.49
		N	9	9	24	23	23
		Proportion	--	--	.96	.75	.65
	School F	Mean	18.69	17.01	509.53	501.58	512.02
		SD	1.34	1.69		93.97	92.33
		N	18	22	49	43	51
		Proportion	--	--	.87	.82	.79
Consistent PLC scores	School G	Mean	16.81	17.04	496.76	444.67	432.94
		SD	1.94	2.04	89.40	108.57	140.26
		N	9	5	17	9	16
		Proportion	--	--	.82	.54	.48
	School H	Mean	19.82	20.00	480.93	504.79	458.32
		SD	1.08	1.34	91.95	114.35	98.61
		N	8	8	27	24	19
		Proportion	--	--	.81	.80	.79

Unfortunately, the growth in the school as a learning community did not translate into achievement gains for students. To try to gain increased understanding of this school we examined their scores on their literacy assessments in the years between our two study years. From 2007 to 2008, the school average on the literacy assessment dropped by 44 points; and while they rose from 2008 to 2009, the 2009 scores were 20 points below the 2007 levels. More pointedly, the number of students not meeting criteria dropped from .85 in 2007 to .67 in 2008, but only rose to .75 by 2009. Statistically, at no time did the achievement scores and proportion of students meeting criteria differ from those for the district level. However, the lack of statistical differences is an artifact of low power.

Case study set 2 – decline in PLC and achievement. The first school selected for the second set of case studies, School D, was a small school. There were five teachers in the school who responded to the 2006 PLC survey, and 6 in 2009. In 2006, six suspensions had been handed out to five different students for a total of 21 days; most of the offenses were for acts of aggression. The number of students completing the Provincial Assessment in 2007 was 14 and this number rose to 21 by 2009.

School D was located in a small rural community and as a result of declining enrollments the status of the school was reviewed as the school board considered closing the school and transferring students to a town 30 minutes away. School D was a very old structure and needed several million dollars to renovate it to keep it open. Morale of school staff members and the school community might have suffered as the school fought to maintain its existence. Community members argued that if their school closed the community would go out of existence. The low enrollment levels also meant that much of the schooling offered at the elementary level was offered in a multi-graded classroom environment, a reality that is often associated with lowered achievement.

Teachers in School D had a positive view of their school as a professional learning community. Their mean PLC score was higher than the district average ($t=3.83, p=.02$). However, not only were the teachers more positive, they were consistent in their opinion. The variance in the teachers' scores was much smaller than the district level variance, indicating uniformity of response ($F=10.34, p=.03$). By 2009, the teachers' views had diminished somewhat, and the mean of their PLC scores had declined to be below that of the district average ($t=6.12, p=.0007$). As in 2006, the variance of their scores was small, indicating uniformity in their responses ($F=27.18, p=.01$).

Unfortunately, the decline in teachers' perceptions of their school as a professional learning community was accompanied by a decline in student achievement. Over the period from 2007 to 2009, student reading scores on the Provincial Assessment declined to the point that in 2009 they were statistically different from the district average ($t=2.80, p=.005$). Given that the reading scores are normed to the provincial average and the low statistical power resulting from small sample sizes, the statistically detectable difference is significant. Concomitant with the decline in scores was a decline in the proportion of students meeting the minimum criteria decreased. By 2009, the proportion of students in School D meeting criteria had declined to fewer than two-thirds and this was statistically lower than the district rate ($\chi^2=4.92, p=.03$).

School E, the next case study in set two, had nine teachers complete the PLC survey in 2006 and 2009. Although the school seems small, in 2006 the school passed out 41 suspensions to 11 students for a total of 60 days. One student accounted for 36% of the suspensions, and three students accounted for nearly 75% of them. In 2007, 24 students completed the Provincial Assessment and this number remained constant over the three-year period.

Although the mean of the teachers' PLC scores was not statistically different from the district level in both 2006 and 2009 (an artifact of low statistical power), the drop in scores was greater than twice the standard deviation of the change in scores over time. The decrease in teachers' perceptions of their school as a professional learning community was accompanied by a steady decline in achievement scores. In 2007, students in School E scored .25 standard deviations above the average and nearly all of the students in Grade 6 met the minimum criteria. Over the year, the school's average on the achievement test dropped approximately 75 points to a level that was nearly statistically different from the district level ($t=1.88, p=.06$) and the proportion of students meeting the minimum criteria had decreased to .75. The decline continued over the next year to the point where the school's scores on the Provincial Assessment were statistically lower than the district level. The number of students achieving criteria had declined to less than two-thirds which approached the statistically detectable level ($\chi^2=3.38, p=.07$).

School F was a relatively large school with 18 teachers completing the PCL survey in 2006, and this number grew to 22 by 2009. There were 49 grade six students who completed the assessment in 2006, and this number was fairly constant over the three-year period. In 2006, four suspensions had been handed out to three different students for minor offenses. This school was located in an affluent community that had strong school board representation and a relatively stable population. Their school underwent an expensive renovation and had a strong principal who worked hard to increase student learning in the school.

In 2007, the average PLC score in School F was similar to the district level ($t=1.02, p=.31$). However, the variance of the scores in School F was less than that of the district to a degree that approached statistical significance ($F=2.17, p=.06$), suggesting a trend towards consensus by the teachers in the school. By 2009, the average score had dropped to a level that was statistically lower than the district average ($t=2.89,$

$p=.004$), and the variance had increased slightly to be statistically similar to that of the district.

If the teachers' perceptions of their school as a professional learning community decreased over a three-year period, it did not show up in the students' achievement levels. Over the three-year period, the scores in School F did not differ statistically from the district level, and showed little variation over time. This might have been the result of the influence of a capable principal who was a strong leader but was resistant to adopting learning community practices within the school. The school principal was confident in his/her leadership ability and sometimes challenged board leaders to provide evidence that this approach would result in increased student learning. The principal was a strong advocate on behalf of the school and was successful in obtaining substantial improvements to the school building and its educational resources.

Case study set 3 – stable PLC and achievement. Two schools were chosen for the third set of case studies, one with relatively low PLC scores and the other with fairly high PLC scores. School G was a relatively small school with nine teachers completing the PLC survey in 2006, and by 2009 this number had decreased to five. There were 17 grade six students who completed the assessment in 2006; one year later this had dropped to a low of nine, but rose back up to 16 in 2009. In 2006, the school handed out 22 suspensions to twelve students, with two students accounting for nearly one-half of them. Most of the suspensions were for minor offenses, such as being disrespectful.

School G experienced some staff problems during our study period and there were other problems that were related to issues of race that might have been affecting within school conditions. The district was aware of the difficulties within the school and provided extra staff allocations to overcome the problems that had developed. The school itself was one of the oldest in the province but was very well maintained and resourced. Even though the school was small, the Board provided a full extra teacher allocation to try to overcome the perceived staffing difficulty.

In 2006, the mean PLC score was statistically below the district level ($t=2.18$, $p=.03$). In 2009 the average had inched upwards enough to be statistically similar to the district mean; however, the change in scores was less than .15 of a standard deviation of slopes across schools. Achievement scores, though, plummeted over that time period. In 2007, the students in School G were achieving close to the district and provincial averages. However, in 2008, the scores had fallen approximately 50 points and by 2009 the school average was more than 60 points (or .6 standard deviations) below the 2006 level ($t=1.97$, $p=.07$). Similarly, the proportion of students meeting criteria fell to just over one-half in 2008 ($\chi^2=5.38$, $p=.02$) and just under one-half in 2009 ($\chi^2=14.44$, $p=.0001$). Of particular note is that the variance in achievement scores for 2009 was larger than that for the district ($F=1.88$, $p=.04$) and suggests that students in School G either did well or poorly with a considerable number doing poorly.

The mean PLC score in School G was slightly but statistically higher than the district average in 2006 ($t=4.18$, $p=.004$) and 2009 ($t=3.74$, $p=.007$). Yet, the change in scores was less than .15 of a standard deviation of slopes across schools. In 2006, the variance of the scores in School G was smaller than that for the district and approached statistical significance ($F=3.35$, $p=.09$) indicating a tendency towards uniformity in responses. Student achievement levels in 2007 were somewhat below the district average, although this was not a statistically detectable difference ($t=.79$, $p=.43$). In 2008, the students in School G were at a level comparable to the district, but in 2009 the average score for the school dropped to a level below the district average that approached statistical significance ($t=1.84$, $p=.07$).

The last school examined in the case studies, School H, was a relatively small school. Eight teachers responded to the PLC survey in both 2006 and 2009. In 2007, there were 27 grade six students who completed the Provincial Assessment and this had decreased to 19 by 2009. In 2006, the school had issued 59 suspensions to 22 students for a total of 149 days. The reasons for the suspensions ranged from disrespect to physical aggression and vandalism. School H was located in an area near a large corrections facility. Inmate families frequently moved to the school catchment area to be near incarcerated family members. This instability among the student population meant that variances in student scores might not be related to school conditions as much as to changes in the composition of the student body. The school principal was widely regarded as an effective leader who experienced high levels of success in other schools.

The mean PLC score in School H was slightly but statistically higher than the district average in 2006 ($t=4.18$, $p=.004$) and 2009 ($t=3.74$, $p=.007$). Yet, the change in scores was less than .15 of a standard deviation of slopes across schools. In 2006, the variance of the scores in School H was smaller than that for the district and approached statistical significance ($F=3.35$, $p=.09$) indicating a tendency towards uniformity

in responses. Student achievement levels in 2007 were somewhat below the district average, although this was not a statistically detectable difference ($t=.79, p=.43$). In 2008, the students in School H were at a level comparable to the district, but in 2009 the average score for the school dropped to a level below the district average that approached statistical significance ($t=1.84, p=.07$).

Summary of findings from case studies. The case studies reveal some interesting findings about schools as professional learning communities and students reading achievement. School A and School C had nearly identical PLC scores in 2006 ($t=.04, p=.97$) and exhibited similar growth with 2009 PLC scores that were not statistically different ($t=.47, p=.64$). Yet the schools displayed different achievement patterns. School A exhibited a 40-point increase in reading achievement relative to the rest of the Province. School C experienced fluctuations in reading achievement. Although these fluctuations were not statistically different from the district average, there was a 44-point drop from 2007 to 2008. At the least, growth in PLC practices did not result in any gains in reading achievement.

School D and School E had similar PLC scores in 2006 ($t=.20, p=.85$) and 2009 ($t=.62, p=.55$). Declines in PLC scores over the three-year period were matched by declines in reading achievement to the point where they were statistically different from the district average. Given the lower power of the statistical tests and the fact that the scores are relative to the rest of the Province, this decline is significant. However, teachers in School F reported a decline in PLC practices but the reading levels in this school did not vary over time.

Discussion

This particular school district invested heavily in initiatives designed to develop and implement Professional Learning Community practices. Our analyses suggest that the investment yielded mixed results. The analyses of changes in teachers' perceptions showed, on average, minimal change over the three-year period. This may be due, in part, to the fact that the average score was above the mid-point of the scale. That is, most teachers had a positive view of their school as a Professional Learning Community. However, there was considerable variation in teachers' perceptions. Some teachers reported considerable improvements in PLC practices while others reported declines, and these changes were reflected at the school level. That is, over time there were discernible school effects that suggested teachers' perceptions of their schools as engaging in PLC practices varied across schools.

At the same time, it was common to see wild fluctuations in achievement scores within schools. A school with low achievement scores in one year might have high scores in the next. Schools might experience achievement declines only to be followed by growth. These fluctuations suggest that performance may be as much a function of individual students as from leadership initiatives.

Despite the mixed results of the present study, Canadian students and students from the study area continue to score very well on international comparative assessments. The wish of Canadian principals to want their teachers to increasingly share in leadership within their schools might be contributing to the enhanced success of Canadian students in comparative international measures. Stronge, Richard, and Catano (2008) contend that guiding a staff to reach a common vision requires intensive and sustained collaboration; this can be best achieved in an environment where teachers are brought into the decision making process. Schools in the study board where our data were gathered have been using research approaches such as those proposed by Fullan (1991, 1993, 1999, 2001), Hargreaves and Fink (2003) and Lieberman and Friedrich (2007) for most of the past decade, but the Board identified it as one of its three goals in 2006.

An earlier study by Hurley (2007) provided evidence that school leaders in the study area were favourably disposed to sharing decisions, and leading in a collaborative manner; these leaders had high expectations that their teachers would take on such responsibilities. Seashore-Louis (2007) suggested that this was a possible spinoff of the political structure of teacher unions in Eastern Canada. She elaborated that having teachers, principals, and many of their supervisors in the same union likely had an effect of encouraging such collegial practices. Another explanation could be that Canadian principals sampled are secure in their leadership positions and perceive that sharing their decision-making might possibly strengthen their leadership roles rather than weaken them. PLC proponents identified these practices as PLC practices. Others in the area of organizational learning (e.g., Leithwood et al., 2004; Senge et al.,

2000; Sheppard et al., 2009) had earlier identified such practices as being linked to a branch of study known as organizational learning. While an apparent overlap of these practices surely exists, we wanted to determine if those particular practices that are described as PLC practices can account for significant variances in student learning. The findings are not conclusive.

In spite of a heavy emphasis by school board leaders and the collaborative development, introduction, and implementation of PLC practices before and during the study period, no significant measurable effect on student learning was identified. This particular study reports only the relationship between PLC practice usage and reading achievement.

It is difficult to draw strong conclusions from the evidence presented here, but there are tendencies that warrant further investigation. First, there is tendency for increases in PLC to be accompanied by increased achievement. Alternatively, if reading levels are average or high then they might be stable. Second, if PLC scores decline, there is a good chance that there will be a decline in reading achievement as well, but this is not necessarily the case. If PLC scores are low but constant, achievement may decrease. If they are high and constant, achievement may be somewhat stable.

References

- Anderson, S., Moore, S., & Sun, J. (2009). Positioning the principal in patterns of school leadership distribution. In K. Leithwood, B. Mascall, & T. Strauss, (Eds.). *Distributed leadership according to the evidence*. New York, NY: Routledge.
- Atlantic Canada English Language Arts Curriculum Guide: Grades 4-6. Fredericton: New Brunswick Department of Education.
- Crowther, F., Kaagan, S., Ferguson, M., & Hann, L. (2009). *Developing teacher leaders: How teacher leadership enhances school success*. Thousand Oaks, CA: Corwin.
- Dufour, R. (2004). Leadership is an affair of the heart. *Journal of Staff Development*, 25(1), 67–68.
- Dufour, R., Eaker, R., & Dufour, R. (2005). *On common ground: The power of professional learning communities*. Bloomington, IN: Solution Tree.
- Elmore, R. (2000). *Building a new structure for school leadership*. Washington, DC: The Albert Shanker Institute.
- Fullan, M. (1991). *The new meaning of educational change*. New York, NY: Teachers' College Press.
- Fullan, M. (1993). *Change forces: Probing the depths of educational reform*. London, UK: Falmer Press.
- Fullan, M. (1999). *Change forces: The sequel*. London, UK: Falmer Press.
- Fullan, M. (2001). *Leading in a culture of change*. San Francisco, CA: Jossey-Bass.
- Garson, D. (2009). Structural Equation Modeling. Retrieved from <http://faculty.chass.ncsu.edu/garson/pa765/>
- Giles, C. & Hargreaves, A. (2006). The sustainability of innovative schools as learning organizations and professional learning communities during standardized reform. *Educational Administration Quarterly*, 42(1), 124–156.
- Hall, G. & Hord, S. (2006). *Implementing change: Patterns, principles, and potholes*. Toronto, ON: Pearson Education.
- Hargreaves, A. & Fink, D. (2003, May). Sustaining leadership. *Phi Delta Kappan*, 84(9), 693-700.
- Hu, L. & Bentler, P. (2000). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Hurley, N. (2007). *Professional learning communities: An effective approach to student learning*. A presentation at the Annual General Meeting of the American Educational Research Association, Chicago, IL.
- Leithwood, K. & Mascall, B. (2008). Collective leadership effects on student achievement. *Educational Administration Quarterly*, 44(4), 529-561.
- Leithwood, K., Seashore-Louis, K., Anderson, S., & Wahlstrom, K. (2004). *How leadership affects student learning*. St. Paul, MO: Learning From Research Project: University of Minnesota, Center for Applied Research and Educational Improvement.
- Lieberman, A. & Friedrich, L. (2007, September). Teachers, writers, leaders. *Educational Leadership*,

65(1), 42-47.

- Marzano, R., Waters, T., & McNulty, B. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Mascall, B., Leithwood, K., Strauss, T., & Sacks, R. (2009). The relationship between distributed leadership and teachers' academic optimism. In A Harris (Ed.). *Distributed leadership* (pp. 81-100). Dordrecht, Netherlands: Springer Science+Business Media B.V.
- Menlo, A. (2011, April). *Responsibility taking among teachers in Michigan*. A presentation to the Consortium for Cross Cultural Research in Education at the Annual General Meeting of the American Educational Research Association, New Orleans, LA.
- Muthén, L.K. & Muthén, B.O. (1998-2012). *Mplus User's Guide*. (7th ed.). Los Angeles, CA: Muthen & Muthen.
- Raudenbush, S. & Bryk, A. (2002). *Hierarchical linear models* (2nd ed.). London, UK: Sage.
- Reeves, D. (2008). *Reframing teacher leadership to improve your school*. Alexandria, VA: Association for Supervision and Curriculum Development.
- SAS Institute. (2010). *SAS/ETS(R) 9.2 User's Guide*. Retrieved from http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/viewer.htm#etsug_model_sect042.htm
- Seashore-Louis, K. (2007). *A discussion of the cross-country study of shared leadership in ten countries*. A presentation to the Consortium for Cross Cultural Research in Education at the Annual General Meeting of the American Educational Research Association, Chicago.
- Senge, P., Cambron-McCabe, N., Lucas, T., Smith, B., Dutton, J., & Kleiner, A. (2000). *Schools that learn*. New York, NY: Doubleday.
- Sheppard, B. & Brown, J. (2000). The transformation of secondary schools into learning organizations. In K. Leithwood (Ed.). *Understanding schools as intelligent systems* pp. 293-314). Greenwich, CT: JAI Press.
- Sheppard, B., Brown, J., & Dibbon, D. (2009). *School district leadership matters*. New York, NY: Springer Science+Business.
- Sheppard, B. & Dibbon, D. (2011). Improving the capacity of school system leaders and teachers to design productive learning environments. *Leadership and Policy in Schools*, 10, 1-21.
- Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models and individual growth models. *Journal of Behavioral Statistics*, 24, 323-355.
- Stronge, J., Richard, H., & Catano, N. (2008). *Qualities of Effective Principals*. Alexandria, VA: Association for Supervision and Curriculum Development.
- York-Barr, J. & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research*, 74(3), 255-316.