

**ADMINISTRATORS' VIEWS ON TEACHER EVALUATION:
EXAMINING ONTARIO'S TEACHER PERFORMANCE APPRAISAL**

[Sachin Maharaj](#), *OISE, University of Toronto*

This study examines the views of administrators (i.e., principals and vice-principals) in Ontario, Canada, with regard to the province's Teacher Performance Appraisal process. A total of 178 responses were collected from a survey that examined five areas: 1) preparation and training; 2) classroom observations; 3) preparing the formal evaluation; 4) the impact on teaching practice; and 5) improving the process. Results indicate that administrators did not receive extensive training and, of the training they did receive, most did not find it very useful. Most administrators did not feel strongly that the classroom observations adequately assessed teacher practice and most did not feel that there had been substantial improvement in teacher practice in their schools as a result of the process. The most common suggestions for improvement were to have more classroom observations, some of which are unannounced; to evaluate teachers more frequently; and to have more than two rating categories.

Introduction

Many research studies have shown that the quality of instruction is the single most important factor in student achievement (Clotfelter, Ladd, & Vigor, 2007; Ferguson & Ladd, 1996; Hattie, 2002; Haycock, 1998; Nye, Konstantopoulos, & Hedges, 2004; Rice, 2003; Rivkin, Hanushek, & Kain, 2005; Whitehurst, 2002; Sanders & Rivers, 1996). The evidence also indicates that having above average teachers for a sustained period of time can overcome the achievement gap between students from higher income and lower income families (Bracey, 2004; Rivkin, Hanushek, & Kain, 2001).

All of this indicates that if we are to close the achievement gap we must ensure that all students, but especially our neediest, have the most effective teachers possible. Every day that

children are subjected to an ineffective or mediocre teacher allows the achievement gap to widen and we know that good teaching has the potential to close it. It should thus be clear what the principal's most important job is (or should be): to ensure high quality teaching in every classroom (Marshall, 2009).

But how does the principal do this? Some suggest that the solution lies in hiring and firing practices. This view was summarized in Newsweek's March 2010 cover story "Why We Must Fire Bad Teachers" addressing the problems of America's education system: "Nothing, then, is more important than hiring good teachers and firing bad ones" (Thomas & Wingert, 2010). The problem with this view is that, within the confines of current collective bargaining agreements, the vast majority of teachers are not candidates for dismissal and, in a climate of declining enrolment (in Ontario as well as in other provinces), vacancies are occurring less frequently. So while more effective hiring and firing practices may increase average teacher effectiveness over time, it fails to address the majority of teachers who are currently in classrooms.

The response of many education systems across the world has been to assess and evaluate its teachers. Teacher assessment and evaluation policies exist in many forms across school districts in the United States, Europe, Asia, and Latin America (Larsen, 2009). They have received increased attention as part of the overall move towards greater accountability in education and efforts to engender a high quality teaching profession. This study aims to examine one such teacher assessment and evaluation policy from the perspective of administrators, the Teacher Performance Appraisal (TPA) process which was introduced to the province of Ontario in 2002.

Significance of the Study

Previous research on teacher assessment and evaluation policies has focused on teachers' perceptions of these policies and has been conducted primarily in the United States (e.g. Johnson, 1990; Kauchak, Peterson, & Driscoll, 1985; Peterson, 2000; Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984). They have largely documented teachers' dissatisfaction with the policies due to perceptions that the administrators conducting the evaluations lacked the resolve, competence, or knowledge to evaluate them effectively and did not provide useful feedback.

With regard to Ontario, although not extensive, there has been some research conducted on teacher assessment and evaluation in the province. Prior to the introduction of the TPA in 2002, there were differing policies with regard to teacher evaluation in school boards across Ontario. In 1988, Lawton, Hickcox, Leithwood, & Musella examined the performance appraisal systems in 30 school boards across the province and found that almost all teachers (over 80%) saw little or no value in terms of professional improvement resulting from the appraisal process. Since the introduction of the TPA to Ontario in 2002, there have been a few studies examining teacher perceptions of the process. The studies have focused on whether teachers feel the TPA process is being applied in a discriminatory fashion (Miller, 2009), whether they feel it accurately assesses their performance (Barnett, 2006), and whether the policy is being applied consistently to teachers across the province (Larsen, 2009).

The problem is that while these, like most other studies on teacher assessment and evaluation, were conducted from the perspective of teachers, minimal research attention has been directed to the perspectives of those who are actually tasked with conducting the process:

administrators. The current lack of focus on the perspectives of administrators has been recognized as a serious deficiency in the literature. As Larsen (2009, p.3) states:

Focusing on teachers' perspectives and attitudes does not preclude the need for further research on performance appraisals from the perspective of vice/principals. . . . More research is clearly needed in this area given the key role that vice/principals play in TPAs and current shifts to implement performance appraisals for principals and vice principals in Ontario. (p. 3)

This study aims to address this deficiency by examining the perspectives of administrators across Ontario in conducting the TPA process. While an individual teacher in Ontario is evaluated only once every five years, administrators in every school across the province spend considerable amounts of time each and every year evaluating the teachers in their schools. The experience of administrators should therefore be of interest to both researchers and policymakers alike. This will allow us to have a more complete understanding of how the TPA process has been implemented in Ontario's schools and the extent to which it has been effective in assessing and improving teacher practice.

Research Question

The study aims to address the following primary research question:

- What are administrators' (i.e., principals' and vice-principals') perceptions regarding the effectiveness of Ontario's TPA process in assessing and improving teacher practice?

Within this primary research question there are seven sub-questions:

1. What are administrators' perceptions regarding the adequacy of the preparation and training they are being provided in conducting TPAs?
2. What sources of information do administrators consult when evaluating teachers?

3. To what extent does subject area expertise play a role in conducting accurate evaluations?
4. How does their role as an evaluator affect administrators' relations with their teaching staff?
5. What are administrators' perceptions regarding the accuracy of the TPA process in assessing teacher practice?
6. To what degree do administrators feel that the TPA process has improved teaching practice in their school?
7. How do administrators feel the TPA process can be improved?

History of Teacher Evaluation

Teacher appraisals have existed in many forms since the introduction of publicly funded schooling in North America. In seventeenth and eighteenth century America, evaluation was used sporadically and primarily as a means of dismissing teachers for blatant incompetence (Vold, 1985). This usually involved a serious breach of either professional responsibilities or the strict moral code of the times.

Scholarly research on teacher evaluation began in 1905 with the work of J. L. Meriam. Meriam demonstrated that there was a low correlation between the grades obtained in teacher-training programs, scholarship in those programs, and actual teaching ability (Peterson, 1982). Later, Barr and Burton (1926) examined teacher appraisal policies in schools and found that “rating schemes force the teacher to live up to the scheme and not to teach good pedagogical principles. Rating forces the teacher to play to the rater and not to the children's interests” (p. 68). Morris (1930) reported on the appraisal policies of many school districts in the United States as it related to teacher pay. Districts across the country used different methods and procedures in evaluating teachers, a trend that has seemingly continued to this day.

In Ontario during this time, school inspectors were primarily concerned with the enforcement of autocratic rules rather than any real beneficial supervision (Houston & Prentice, 1988; Prentice & Theobald, 1991). An example of these enforced rules was whether or not the schedule regulating the exact amount of time to be spent on each subject was being followed precisely (Danylewicz & Prentice, 1986). Overall, this period of teacher evaluation in Ontario could be characterized as one that stressed strict uniformity and unquestioned obedience (Salisbury, 1912).

American Literature on Teacher Evaluation

Since then, many studies have been conducted on teacher appraisals but most have been conducted in the United States and from the perspectives of teachers. Peterson's (2000) extensive literature review of over 70 years of empirical research on teacher evaluation concluded: "Seventy years of empirical research on teacher evaluation shows that current practices do not improve teachers or accurately tell what happens in classrooms. . . . Well designed empirical studies depict principals as inaccurate raters both of individual teacher performance behaviours and of overall teacher merit" (pp. 18–19).

During this time, a plethora of interview survey studies were conducted on teachers' perceptions of their evaluations. The results were almost uniformly negative. A representative sample of five of these studies will be presented. In 1973, Wolf reported an interview study with 293 teachers. In general, the study found that teachers mistrust evaluation:

They feel that current appraisal techniques fall short of collecting information that accurately characterizes their performance. They perceive the ultimate rating as depending more on the idiosyncrasies of the rater than on their own behaviour in the classroom. As a result, teachers see nothing to be gained from evaluation. (p. 160)

Lortie (1975) found that only 7% of the teachers he interviewed saw judgements by their organizational superiors as the most appropriate source of information about how well they were doing. The study concluded that teachers had little direct interest in or respect for the process or results of evaluation, and most operated independently of them.

Kauchak, Peterson, and Driscoll (1985), in a survey study of teachers in Utah and Florida, found evaluations based on principal visits to be “perfunctory with little or no effect on actual teaching practice” (p. 33). One problem identified by the teachers in the study was that evaluations were too brief and lacked rigour. Teachers also complained that the principal was not knowledgeable in their grade level or subject area. Finally, teachers in the study felt that the evaluation reports lacked specifics about how to improve their teaching practice.

Wise, Darling-Hammond, McLaughlin, and Bernstein (1984) completed an extensive survey and case studies on teacher evaluation for the National Institute of Education. The most significant problem they found was that principals were being put in a significant role conflict position:

Central office respondents believed that the conflict between principal as instructional leader and evaluator has not been settled. Noting that collegial relationships lead many principals to want to be “good guys,” many respondents felt that principal evaluations were upwardly biased. Principals’ disinclinations to be tough makes the early identification of problem teachers difficult and masks important variations in teacher performance. (p. 22)

The study also found that principals considered teacher evaluation “a necessary evil or time-consuming chore” (p. 22).

In addition, four other major problems with teacher evaluation were identified in the study by Wise et al. (1984). One was teacher apathy, as full teacher support was reported in less than half the districts surveyed. Second, was lack of uniformity and consistency within school districts. Teachers reported that evaluations depended too much on the predisposition of the

principal and that similar teacher practices led to different ratings in different schools. Third, was the inadequate amount of training given to principals in conducting evaluations. Central office respondents reported that principals received too little training and guidance for their current evaluation responsibilities. The final problem was expertise. Many teachers, especially high school teachers and specialists felt that many administrators lacked the subject matter knowledge necessary for an accurate evaluation.

Johnson (1990) interviewed 115 teachers and found similar results. Teachers felt that principals rarely offered ideas for improvement. They also felt that the ratings forms and items encouraged principals to be picky in their criticisms; almost forcing principals to find something to criticize so that they will look discriminating. However, the main dissatisfaction of teachers in the study was what teachers saw as a basic lack of competence on the part of administrators to evaluate. This included a lack of self-confidence, expertise, subject matter knowledge, and perspective on what it is really like to be in the classroom.

Odden (2011) has noted more recently that teacher evaluation in the United States is generally of low quality yet consumes a great deal of time for both teachers and administrators. Furthermore, it is rarely used for anything other than terminating the worst performers and compliance with government regulations and is therefore not strategic (Weisberg, Sexton, Mulhern, & Keeling, 2009).

The American literature on teacher evaluation indicates that neither teachers nor administrators seem to receive much benefit from the process, despite it consuming large quantities of time and resulting in considerable stress. The impact on teaching practice appears to be negligible and often results in negative feelings among teachers as they do not feel that their

evaluations are objective or accurate. Administrators often view teacher evaluations as something they are forced to do rather than something they want to do.

Ontario's TPA Policy—Background

Prior to 2002, there was no uniform system of teacher appraisal in school boards across Ontario. In fact, a study conducted in 1982 found that over a quarter of all school boards in Ontario had no teacher appraisal policy whatsoever (Lawton, Hickcox, Leithwood & Musella, 1988). Where policies did exist, they consisted of differing elements that were implemented in inconsistent ways across school boards. For example in some boards teacher evaluation was conducted solely by principals; in others, superintendents evaluated teachers as well. In the secondary schools of some school boards, department heads were also involved in the process of teacher evaluation. With regard to the actual process of evaluation, 20% of teachers in Ontario reported having three or more classroom observations per evaluation whereas more than 50% reported having one or none at all.

This inconsistency led to calls by the government of Mike Harris for a uniform system of evaluation across the province. As the Minister of Education, Janet Ecker, stated prior to the introduction of the TPA: “The Act to Promote Quality in the Classroom, 2001, will, if passed, provide for fair and consistent standards for teacher performance appraisals in every school” (Ecker, 2001). In December 2001, the Harris government successfully amended the Education Act through the Quality in the Classroom Act. After the Act had passed, the Ministry of Education (2002) stated three specific purposes for the TPA:

- to ensure that students receive the benefit of an education system staffed by teachers who are performing their duties satisfactorily;

- to provide for fair, effective, and consistent teacher evaluation in every school;
- to promote professional growth. (p. 3)

In addition to the introduction of the TPA the Act mandated Annual Learning Plans (ALPs) for all teachers. By the fall of 2002, all teachers in the province were to have received the TPA manual, which outlines the entire TPA process, as well as their ALP form. The entire TPA system was to be fully implemented in all schools by 2004. Experienced teachers were to be evaluated every three years and new teachers were to be evaluated twice in each of their first and second years of employment. For experienced teachers, two evaluations were required during their evaluation year.

The TPA policy was changed by the McGuinty government in 2007 and this is the current policy in place at the time of this study. The TPA now has four specific purposes:

- promote teacher development;
- provide meaningful appraisals of teachers' performance that encourage professional learning and growth;
- identify opportunities for additional support where required; and
- provide a measure of accountability to the public. (Ministry of Education, 2010, p. 5)

Instead of being evaluated every three years, experienced teachers are now evaluated every five years. The requirement of two evaluations in the evaluation year has been reduced to one evaluation. New teachers are evaluated twice in only their first year of teaching, as opposed to be evaluating twice in each of their first two years under the old system. Besides the frequency of evaluations, the biggest change was in the ratings given to teachers. Under the old system, teachers were given one of four ratings: unsatisfactory, satisfactory, good, or exemplary. The

new system eliminated the last two categories and now teachers receive one of two ratings: unsatisfactory or satisfactory.

There are two main reasons why the McGuinty government changed the Teacher Performance Appraisal policy. First, under the original system of multiple rating categories and more frequent evaluations, many administrators complained of spending an inordinate amount of time both evaluating their teachers and dealing with teacher union officials, as many teachers challenged the ratings they received. Thus, evaluations that were conducted less frequently and with a system of only two rating categories were a way to address this issue. But, second and probably more significantly, the election of the McGuinty government promised a new beginning of relations between the provincial government and Ontario's teachers after the tumultuous tenure of Mike Harris. And mandated teacher evaluations are often opposed by teacher unions, both in the United States and in Canada. For example, the British Columbia Teachers' Federation *2012–2013 Members' Guide* states "the evaluation of members should be based on the assumption of professional competence and, hence, formal evaluation should not occur unless the assumption is questioned or a formal evaluation is requested by the member" (p. 155). Hence a change in the evaluation system that resulted in conditions more agreeable to Ontario's teacher unions was seen as a goodwill gesture to help build trust and good relations between the government and the province's teachers.

In terms of evaluation criteria, each teacher is to be evaluated with respect to the standards outlined in the Ontario College of Teachers *Standards of Practice for the Teaching Profession* (Ministry of Education, 2010). These include commitment to students and student learning, professional knowledge, professional practice, leadership in learning communities, and ongoing professional learning. For the purposes of the TPA, there are 16 competency statements

within the five domains of the Standards of the Teaching Profession (see Table 1). New teachers are evaluated on a subset of 8 of the 16 competencies (see blue highlighted items in Table 1) whereas experienced teachers are evaluated on all 16 competencies.

Table 1
Ontario's TPA Competency Statements

Domain	Competency
Commitment to Pupils and Pupil Learning	<ul style="list-style-type: none"> ▪ Teachers demonstrate commitment to the well-being and development of all pupils. ▪ Teachers are dedicated in their efforts to teach and support pupil learning and achievement. ▪ Teachers treat all pupils equitable and with respect. ▪ Teachers provide an environment for learning that encourages pupils to be problem solvers, decision makers, lifelong learners, and contributing members of a changing society.
Professional Knowledge	<ul style="list-style-type: none"> ▪ Teachers know their subject matter, the Ontario curriculum, and education-related legislation. ▪ Teachers know a variety of effective teaching and assessment practices. ▪ Teachers know a variety of effective classroom management practices. ▪ Teachers know how pupils learn and the factors that influence pupil learning and achievement.

Teaching Practice	<ul style="list-style-type: none"> ▪ Teachers use their professional knowledge and understanding of pupils, curriculum, legislation, teaching practices, and classroom management strategies to promote the learning and achievement of their pupils. ▪ Teachers communicate effectively with pupils, parents, and colleagues. ▪ Teachers conduct ongoing assessment of pupils' progress, evaluate their achievement, and report results to pupils and their parents regularly. ▪ Teachers adapt and refine their teaching practice through continuous learning and reflection, using a variety of sources and resources. ▪ Teachers use appropriate technology in their teaching practices and related professional responsibilities.
Leadership and Community	<ul style="list-style-type: none"> ▪ Teachers collaborate with other teachers and school colleagues to create and sustain learning communities in their classrooms and in their schools. ▪ Teachers work with professionals, parents, and members of the community to enhance pupil learning, pupil achievement, and school programs.
Ongoing Professional Learning	<ul style="list-style-type: none"> ▪ Teachers engage in ongoing professional learning and apply it to improve their teaching practices.

(Ministry of Education, 2010, p. 20)

The primary responsibility for conducting the TPA lies with the principal, although it can be delegated to vice-principals and, in certain circumstances, supervisory officers. Under the original TPA policy, student and parental input were to play a role in teacher evaluation but under the changes brought in by the McGuinty government that requirement has been removed

and it is now merely suggested that teachers use student and parental input when developing their ALP.

The TPA consists of the following elements:

- a pre-observation meeting,
- a classroom observation,
- a post-observation meeting,
- a summative report that includes a rating of the teacher's overall performance. (Ministry of Education, 2010, p. 29)

The pre-observation meeting allows the teacher and administrator to prepare for the classroom observation component. A date is then set at which point the classroom observation occurs. After the classroom observation, the post-observation meeting takes place at which point the teacher and administrator review the results of the observation and discuss other information relevant to the appraisal of the teacher's performance. The post-observation meeting is to be held "as soon as possible after the classroom observation" (Ministry of Education, 2010, p. 31).

Finally, the summative report is prepared. The administrator must prepare and sign the summative report as well as give the teacher a copy within 20 days of the classroom observation. The teacher must then sign the copy to acknowledge receipt and can add comments if he or she desires. The administrator must then send a signed copy to the school board.

Canadian Literature on Teacher Evaluation

There have been only a handful of studies conducted on teacher appraisal in Canada, although, of those, most have been done in Ontario. In the 1980s, commissioned by Ontario's Ministry of Education, Lawton, Hickcox, Leithwood, and Musella (1988) completed a comprehensive study of the different performance appraisal policies (not just of teachers, but at

all levels) in Ontario's school boards. Over 5000 teachers, 1200 principals, 200 superintendents, 150 trustees, and 30 directors were surveyed. Over 80 percent of teachers said they perceived little or no improvement in their teaching performance as a result of the appraisal process. By contrast, over 80 percent of superintendents said they perceived either a modest or substantial improvement in teachers' performance as a result of the appraisal process. Table 2 displays the discrepancy.

Table 2
Degree of Improvement in Teachers' Performance as a Result of Evaluation as Perceived by Evaluatees and Superintendents

Degree of Improvement	Teachers (n=3158)	Superintendents (n=100)
Not at all	40.5%	0.0%
A small amount	43.6%	14.0%
A modest amount	13.3%	43.0%
A substantial amount	2.6%	43.0%

(Lawton, Hickcox, Leithwood, & Musella, 1988, p. 32)

Although the study included administrators, it mainly focused on their perceptions of their own appraisals by superintendents, as oppose to their appraisals of teachers. However, the study did recognize that the process of conducting evaluations was one that consumed a substantial amount of principals' time: "The pressure placed to collect information about large numbers of people on a regular basis is draining of both physical and psychic energy" (p. 27). Despite it being time consuming, principals felt that more time should be spent on each evaluation. This is a theme that was echoed in more recent examinations of Ontario's TPA process. Bolger and Vail (2003) concluded that there was insufficient time for administrators to

complete the process. Black's (2003) article, although conceptual rather than research based, agreed with those findings.

The Lawton et al. (1988) study also concluded that there was a need for greater training of administrators in the process of conducting evaluations, a finding that was echoed over 15 years later in Bolger and Vail's (2003) study. Bolger and Vail also found that administrators viewed the appraisal process as primarily a competency instrument with professional growth as a by-product.

Rowe (2000) conducted a study on teachers' perceptions of the performance appraisal process in Newfoundland. The teachers in the study expressed dissatisfaction with the hierarchical nature of the appraisal's implementation. This was similar to sentiments expressed by Black (2003) and Bolger and Vail (2003) in Ontario. He concluded that the process was unsupportive of professional growth. However, during the study, the school district altered the performance appraisal process to make professional growth the primary objective and afterwards teachers reported a much better experience.

Prior to the introduction of the uniform TPA to the province of Ontario, Goodman (2001) examined teacher and principal experiences with the teacher evaluation procedures in their various boards. Based on 12 interviews (six principals, six teachers), the study reported a number of findings similar to those of Lawton et al. (1988), Bolger and Vail (2003), and Black (2003). Both the teachers and the principals in the study agreed that the number of visits allotted for classroom observations was inadequate, that there was no clear distinction between the formative (i.e., professional growth) and summative (i.e., competency) components of the evaluation, and that the training provided for conducting the evaluation was inadequate.

After the introduction of the TPA, Barnett (2006) examined Ontario teachers' perceptions of the process as to whether it accurately assessed their performance and encouraged professional growth. The study concluded that teachers felt that their evaluations were not based on their performance but rather their relationship with the administrator and that they did not contribute to professional growth.

Miller (2009) examined the TPA from teachers' perspectives through an equity lens to see if it is being applied in a discriminatory fashion. His mixed methods study, involving six interviews and surveys of 132 teachers concluded that minority teachers tend to experience mistreatment, including racism and homophobia from administrators who conducted the TPA.

Larsen (2009) also examined Ontario's TPA from the perspective of the teacher. Her mixed methods study involving interviews with 25 teachers and a survey of 125 teachers concluded that the majority of teachers feel the TPA process is "disorganized, inconsistently conducted and above all unfair" (p. 24). Teachers reported the process left them with enhanced feelings of stress, anxiety, and self-doubt.

It is worth noting that these studies conducted in Ontario all occurred prior to the revision of the TPA policy in 2007. What should be noted about the Canadian literature more generally is that none of the studies focused on the experience and perceptions of those actually tasked with conducting the evaluation process: administrators. However, this appears to be a common feature of research on teacher evaluation more generally.

The minimal research attention that has been directed to the perspectives of administrators has been recognized as a serious deficiency in the literature. As Larsen (2009) states,

focusing on teachers' perspectives and attitudes does not preclude the need for further research on performance appraisals from the perspective of

vice/principals. . . . More research is clearly needed in this area given the key role that vice/principals play in TPAs and current shifts to implement performance appraisals for principals and vice principals in Ontario. (p. 3)

There thus exists a clear need to focus on the perspectives of administrators when examining Ontario's TPA process. While an individual teacher in Ontario is evaluated only once every five years, administrators in every school across the province spend considerable amounts of time each and every year evaluating the teachers in their schools. The experience of administrators should therefore be of interest to both researchers and policymakers alike. This will allow us to have a more complete understanding of how the TPA process has been implemented in Ontario's schools and the extent to which it has been effective in assessing and improving teacher practice.

Validity of Current Teacher Evaluation Systems

It appears that neither teachers nor administrators view current teacher evaluation systems as being particularly useful. As stated by Odden (2011), "this type of teacher evaluation has been criticized for years as invalid, unhelpful, and largely a waste of teacher and administrator time" (p. 71). From the administrator's perspective it is easy to see why this might be. Evaluations of teacher performance often consist of a single, pre-arranged visit by the principal to observe a teacher's class. As the visit is arranged well in advance, teachers often put on a showcase, one that is not necessarily reflective of their daily teaching practice. While teachers cannot be blamed for wanting to put their best foot forward, no one should be under the illusion that this is an accurate way to provide an evaluation. Here is a true story of one teacher under such an evaluation system:

In preparation for an evaluation visit, a teacher had distributed a special student handout to her class. When she heard over the intercom that the

principal had to postpone his observation, she collected the worksheet from students and proceeded with the “normal” lesson. (Marshall, 2009, p. 23)

Furthermore, current evaluations tend to provide only a snapshot of a teacher's professional practice. If the evaluation consists of a single class visit, and teachers teach an average of four classes per day for 190 days per year, this means that a teacher's evaluation is entirely based on observing 0.1 per cent of their teaching practice. The remaining 99.9% of the time is essentially ignored. A teacher could put on a good show or, conversely, may simply have an off day. In any case, no one lesson can provide a complete picture of teacher practice.

But perhaps the most troubling aspect of current evaluation systems is the subjective nature of the classroom observations upon which they are based. Evidence indicates that different administrators can give quite different ratings when viewing the same lesson (Gates Foundation, 2012). This combined with the previously mentioned finding (Barnett, 2006) that many teachers feel that their evaluations are based more on the nature of their relationship with their administrator, as opposed to their actual teaching practice is troubling. It seriously calls into question whether a teacher's evaluation should be based almost entirely on the subjective opinion of just one person.

Conceptual Framework

In the literature, teacher appraisals have been identified as having two distinct purposes: competency and professional growth commitments (Beerens, 2000; Joyce, Calhoun, & Hopkins, 1999; Middlewood, 2002; Porter, Youngs, & Odden, 2001; Sergiovanni, 1992). Competency commitment is the “systematic investigation of the worth or merit of an object” (Joint Committee on Standards for Educational Evaluation, 1994, p. 3). The object in this case is the teacher, who is evaluated against a set of predetermined standards. Teachers must meet these accepted

standards in order to demonstrate that they are capable in their professional roles. Competency points to teacher cause and effect relationships with respect to instruction, and the use of this information for personnel and accountability purposes such as staffing, tenure, promotion, or dismissal (Rowe, 2000).

The other purpose of teacher appraisals is to enhance professional growth. This purpose is collegial in nature and teachers assume responsibility for their own learning and professional growth (Cousins, 1995). By emphasizing reflection, the professional practice of all teachers can be improved. As Osborne (1987) states,

continual improvement is good for all. It is not enough to simply know who our “good” teachers are; we must work for the improvement for all of our staff members—both “good” and “poor.” (p. 7)

This framework of competency and professional growth aids this research study in addressing the two key areas of the main research question: What are administrators' perceptions regarding the effectiveness of Ontario's TPA process in assessing (*competency*) and improving (*professional growth*) teacher practice? As several research studies have concluded that these purposes are often in conflict (Beerens, 2000; Joyce et al., 1999, Middlewood, 2002; Porter, Youngs & Odden, 2001; Sergiovanni, 1992), it will be an important contribution to the body of knowledge to explore to what degree administrators perceive they are able to accomplish both purposes when conducting the TPA process.

In terms of the actual process of conducting teacher appraisals, Lawton, Hickcox, Leithwood, and Musella (1988) have constructed a framework which identified four distinct aspects: preparation for appraisal, data collection, reporting and follow up, and impact and evaluation. Preparation for appraisal includes training in conducting evaluations, planning activities by supervisors, establishment of objectives, agreement on criteria and the establishment

of standards by which it can be determined if the objectives have been met. Data collection is concerned with the sources of information on which an evaluation is based, the type of information collected, who collects the information, and the time and effort spent collecting the information. Reporting and follow up include post-observation conferencing and the preparation of formal evaluation reports. Impact and evaluation is the examination of the degree to which the teacher appraisal process was successful in meeting its objectives. This study aims to examine administrator experiences in all four domains.

The Figure demonstrates the conceptual framework for this study. Table 3 outlines the aspects that will be examined in each domain. By examining the experiences of administrators in each of these domains, this study aims to assess the degree to which administrators perceive the current TPA process is effective in achieving the two objectives of competency and professional growth (shaded in green).

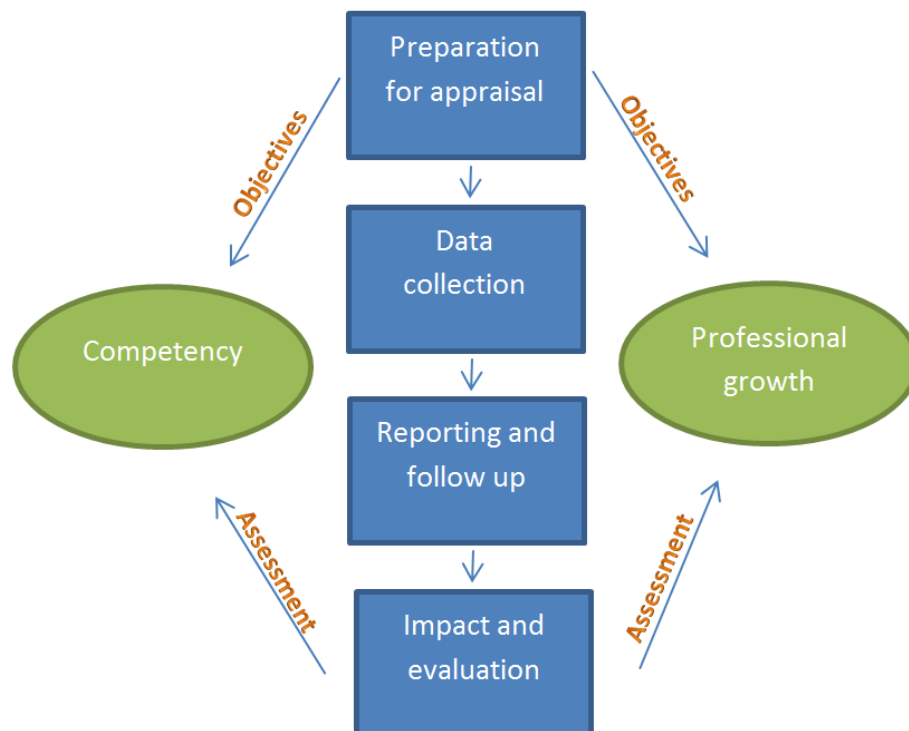


Figure. Conceptual framework.

Table 3
Variables Being Examined in Each Domain of the Teacher Appraisal Process

Domain	Description
Preparation for appraisal	<ul style="list-style-type: none">▪ Training▪ Planning activities by supervisors▪ Establishment of objectives
Data collection	<ul style="list-style-type: none">▪ Sources of information on which evaluation is based▪ Time and effort spent collecting the information
Reporting and follow up	<ul style="list-style-type: none">▪ Post observation conferencing▪ Preparation of the formal report
Impact and evaluation	<ul style="list-style-type: none">▪ Assessing effectiveness in meeting objectives

Methodology

The instrument for this study was an electronic web-based survey. The survey questions (see Appendix) were structured around the four domains and two objectives of teacher appraisal that were outlined in the conceptual framework. The survey utilized a combination of personal, attitudinal, and behavioural questions. In terms of question format, a combination of open-ended, closed-ended, and semi-closed ended questions were used. Semi-closed ended questions were the most common format as it allows for the advantages of both open and closed-ended questions.

Data Collection

The study was conducted with the support of the Ontario Principals' Council, which represents 5,000 administrators in English-language public schools across Ontario. While this does not include all administrators, such as those in the province's Catholic and French schools,

it does represent the majority of school administrators in Ontario. In the weekly email newsletter that is sent to its members, the OPC provided a description of the study and a link to the informed consent letter and the online survey for interested participants. A total of 178 administrators provided responses. Of those administrators that provided responses, a total of 166 completed the entire survey.

Findings

The following section analyzes the results of the survey through the lens of the conceptual framework for this study. For the full results of the survey, see Appendix.

Preparation for Appraisal

While most administrators reported receiving some training, it does not appear it was extensive as only 34% reported receiving a great deal or a lot of training. When TPAs were first introduced by the province administrators indicated receiving mandatory training, but since then the only training that has been provided has been voluntary workshops put on by school boards and the OPC.

In terms of the content of these voluntary workshops, administrators are being trained in the technical aspects of the TPA, but not in how to accurately assess and evaluate teachers. As one participant put it “I was trained in what all of the aspects of the TPA are. There was not so much information about how to conduct the meetings or observations.” This may explain the fact that while the majority of participants found the training they received at least somewhat useful, only 36% found it extremely or very useful.

Administrators noted that in addition to lacking instruction in how to accurately assess and evaluate teachers, there was also very little training in how to deal with unsatisfactory TPAs. A possible reason given for this is that administrators are being discouraged from giving unsatisfactory TPAs, which would seem to undermine their ability to fulfill the competency objective of teacher evaluations. If administrators are being discouraged from giving negative evaluations, these inaccurate assessments cannot be meaningfully used for personnel or accountability purposes.

Since there does not appear to be extensive formal training provided in conducting TPAs, many administrators report that they review the TPA manual and other policy related documents prior to conducting their appraisal. They also meet with teachers to review the process and requirements of the TPA and establish timelines for classroom observation and conferences.

Data Collection

Administrators reported spending an average of 86 minutes observing teachers during the classroom observation portion of the TPA. Administrators did not feel strongly that the classroom observations were adequate in accurately assessing teacher practice, with only 26% stating that they were extremely or very adequate. Many administrators felt that the pre-planned nature of the classroom observations did not allow for an accurate and thorough assessment of a teacher's regular teaching practice. As one administrator put it, "observing a single class does not give a very thorough picture of a teacher's practice. A teacher can always put on a good show."

Thus it is perhaps not a surprise that the majority of administrators felt that more classroom observations would allow for a more accurate and useful teacher appraisal with 64%

stating they strongly agree or agree. However many felt that while more observations would be useful, there was simply not more time in their already busy schedule to allow for this. As one administrator put it, “yes, however where is this time going to come from on the part of the administrator?”

Others felt that the observations needed to be unscheduled in order to provide an accurate depiction of the teacher's daily teaching practice. Administrators reported using a variety of sources other than classroom observations when making their assessment. This included the teacher's lesson plans, unit plans, assessments, notebooks, conduct during staff meetings and professional development sessions, relations with staff and students, and what was often termed “contribution to the school community.”

Interestingly, many administrators also reported utilizing sources of information outside of those strictly dictated by the current TPA process. This included informal visits to the teacher's classes throughout the year (which many administrators mentioned), report card data including the teacher's pass-fail rates, and anecdotal input from curriculum leaders or department heads. It appears that administrators may be using these “outside” sources of information because they feel that the current sources of information available to them within the guidelines of the TPA process (e.g., scheduled classroom observations) are inadequate.

Despite the fact that many administrators commented that subject area expertise is not required in order to effectively evaluate a teacher, administrators reported being much more confident conducting performance appraisals for teachers that taught within their subject area of expertise. Ninety-seven percent of administrators reported being extremely or very confident when evaluating teachers that taught subjects within their area of expertise, compared with 71% that reported being extremely or very confident when evaluating teachers that taught subjects

outside their area of expertise. This is especially relevant as almost all administrators reported having to conduct performance evaluations for teachers that taught subjects outside their own area of expertise and almost half (45%) reported doing so always or most of the time.

Reporting and Follow Up

Administrators reported conducting the post-observation meeting an average (mean) of 4.33 days later, which would appear to meet the ministry requirement that the meeting be held “as soon as possible after the classroom observation.”

Almost all administrators reported rarely having to deal with teachers disagreeing with their assessment with 97% reporting that this happened only once in a while or never. Based on the comments provided, it appears that the only time administrators had teacher disagree with their assessment was when they were given an unsatisfactory rating, which does not appear to happen very often. It is curious whether a possible reason for this is that they are being discouraged from giving such ratings as was previously indicated.

Another possible reason for this is that some administrators reported receiving pressure from union officials when an unsatisfactory rating was given. As one administrator put it, “the federation is a daunting presence in such circumstances. It takes a lot of time and the time is worth it because it is so important. But it is very stressful and the federation can make your life just plain miserable.”

Administrators also reported that there was more disagreement when the previous four point rating scale was in place. As one administrator put it, “this was more of an issue when there was a four point rating scale rather than a two point scale.” Perhaps this could be one reason why the four point scale was scrapped in favour of the current two point scale.

Another possible reason that there was so little disagreement was that once a teacher achieves the satisfactory rating, the TPA is not used for future personnel decisions. As one administrator put it,

I think this speaks to the fact that as long as you are given a satisfactory appraisal, teachers are satisfied. They understand that it matters little. TPAs are never even mentioned when a candidate is applying for a new job with a new school and administrators are looking for a reference. I have never seen a question that asks "How was their last TPA?"

This would appear to undermine the competency objective of teacher evaluations.

Impact and Evaluation

Most administrators did not feel that their role as an evaluator negatively affected relations with their teaching staff, with 83% stating that they disagree or strongly disagree. Most administrators stated that they embraced the evaluation role as part of their job and that in fact it has often enhanced their relations with their teaching staff. Many viewed it as an opportunity to provide coaching and mentorship and to build better relationships with their teachers. Thus it would appear that many administrators embraced the professional growth objective of teacher evaluation. The only instances where it appeared to harm relations were with those teachers that received an unsatisfactory rating, which as previously stated, does not appear to occur often.

Administrators appeared divided on the question of whether the TPA process accurately assesses the practice of teachers with 41% stating they strongly agree or agree, 28% undecided and 32% stating they disagree or strongly disagree. Many felt that the TPA process is only as good as the administrator conducting it. Others noted that the process is good at highlighting the work of great teachers but not at addressing mediocre ones, partly as a result of union involvement. Still others felt the process was too contrived and artificial to be of use.

In terms of professional growth, most administrators felt there had not been substantial improvement of teacher practice in their schools as a result of the TPA process. Just 9% indicated that teacher practice had improved a substantial amount, 34% a modest amount, 36% a small amount, and 20% not at all. This appears to align with the views of the teachers in Ontario surveyed in 1988 by Lawton, Hickcox, Leithwood, and Musella who felt that the teacher evaluation systems at the time did not result in substantial improvement in teacher practice or performance.

Most administrators appeared not to view the TPA process as a major vehicle for improving teacher practice. Instead, many indicated that improvements in teacher practice were the result of ongoing professional development in the school during PLC time, PD days, staff meetings, and coaching and mentoring amongst administrators and staff. Many administrators commented that because there is no follow up for teachers who get a satisfactory rating, improvements in practice usually only occur with those that receive an unsatisfactory rating.

Administrators were also divided on the question of whether conducting TPAs was an effective use of their time, with 47% strongly agreeing or agreeing, 21% undecided, and 32% disagreeing or strongly disagreeing. Many administrators commented that while they enjoyed observing teachers and having an opportunity to engage them about their teaching practice, the time spent on paperwork and other bureaucratic aspects of the evaluation was not as worthwhile. Others felt that without more than two rating categories or follow up for those who receive satisfactory ratings, the process was not as effective as it could be. Yet many still viewed the process as valuable in that it appears to be one of the only vehicles to address poor teaching.

In terms of improving the TPA process, the most common suggestion from administrators was to have more classroom observations. Many administrators felt that basing

the TPA on a single classroom observation was insufficient. In addition, administrators stated that to get a true representation of a teacher's teaching practice at least some of the observations should be unannounced. Administrators also stated that evaluating teachers only once every five years was too infrequent. Many also remarked that there should be more than two rating categories so as to distinguish really excellent teaching from that which is merely satisfactory. Another common suggestion was to streamline the number of competencies that teachers must be evaluated on, as many are very similar. In addition, some administrators also felt that there should be a cap on the number of TPAs they must perform in a given year so that they can give more attention to each one.

Implications for Policy and Practice

While the TPA is certainly an improvement over the inconsistent patchwork of teacher evaluation schemes that existed across Ontario prior to its introduction, the findings of this study indicate that there are many ways that it can be substantially improved so that it better meets its objectives of competency and professional growth.

Administrative Responsibilities

It appears that many administrators would like to spend more time focusing on effective teaching and evaluation but are already overburdened with other administrative tasks. For example, while the majority of administrators agreed that more classroom observations would allow for a more accurate and useful teacher appraisal, many felt that there was simply no more time in their busy schedules to allow for this. As one administrator put it, "yes, however where is this time going to come from on the part of the administrator?"

Thus administrators will need to be relieved of many less critical responsibilities in order for them to adequately meet both the competency and professional growth objectives of teacher evaluation. If ensuring high quality teaching in every classroom is an administrator's most important function, their responsibilities should reflect this. This will also allow administrators to better focus on being instructional leaders in their schools.

Training

Training for administrators in conducting TPAs does not appear to be either very useful or mandatory. Only 34% of administrators reported receiving a great deal or a lot of training in conducting TPAs and only 36% found their training extremely or very useful. Given the importance of teacher evaluation, all administrators should receive regular training in not just the mechanics of the TPA, but also how to accurately assess and evaluate teachers both in and out of the classroom. For new administrators, this training should be mandatory prior to beginning their new position. For experienced administrators, ongoing training should be mandated to ensure that their skills are kept sharp and that there is consistency across schools. There also appears to be a need for more training and support around unsatisfactory TPAs. Given that unsatisfactory TPAs involve substantial work on the part of the administrator, providing more training and support in this area will likely make administrators more confident when dealing with these difficult situations.

Classroom Observations

In the eyes of most administrators, assessing the effectiveness of a teacher based on a single, pre-planned classroom observation is inadequate. Only 26% felt that the classroom

observations were extremely or very adequate in assessing teacher practice. As one administrator put it, "observing a single class does not give a very thorough picture of a teacher's practice. A teacher can always put on a good show." Another administrator expressed similar sentiments: "All observations are staged events which is not necessarily representative of the classroom experience." Thus the TPA should include multiple classroom observations, some of which are unannounced. This would provide administrators with a more accurate picture of a teacher practice and would thus allow for more meaningful assessment and feedback.

Subject Area

Despite comments from some administrators to the contrary, it appears that subject area expertise does matter when conducting teacher evaluations. As evidenced in the literature, this has long been a concern among teachers and it appears that the administrators in this study acknowledge the issue as they reported being much more confident when conducting evaluations for teachers that teach within their area of expertise. Ninety-seven percent of administrators reported being extremely or very confident when evaluating teachers that taught subjects within their area of expertise, compared with 71% that reported being extremely or very confident when evaluating teachers that taught subjects outside their area of expertise. Hence, as much as possible, teachers should be evaluated by administrators that share their area of expertise. While this will not always be feasible, it is certainly preferable as it will likely allow for a more accurate and meaningful evaluation and is more likely to be perceived as such by the teacher being evaluated. This may make teachers more receptive to the feedback they receive as they will have increased confidence in the administrator's evaluative abilities.

Sources of Information

Many administrators appear to be utilizing sources of information outside of those strictly dictated by the current TPA process. This included informal visits to the teacher's classes throughout the year (which many administrators mentioned), report card data including the teacher's pass-fail rates, and anecdotal input from curriculum leaders or department heads. This may be because they feel that the current sources of information available to them are inadequate. One possible way to remedy this is to include other sources of information in the evaluation process. Instead of relying solely on classroom observations or other aspects of teacher behaviour such as the organization of notebooks and unit plans, evaluations could incorporate direct evidence of student learning. This could take many forms including student performance on end-of-course assessments, teacher-generated assessments, district-wide assessments, or value-added models that measure student progress on standardized assessments while controlling for important factors such as students' academic history or demographic characteristics.

Another option is to incorporate feedback from students. This was supposed to be part of Ontario's evaluation process but was removed by the McGuinty government. However, the rationale for its inclusion remains quite valid. Students spend hundreds of hours with their teachers every year and so probably have the most complete picture of the teaching practice and effectiveness of those teachers. With student evaluations of instructors being commonplace in university and college classrooms across the country, there is no reason why this could not also happen in our K–12 system as well. While some teachers may worry that students will simply give the “easy” teachers higher scores and the “hard” teachers lower ones, large scale studies on student surveys of teacher performance show that this is not the case when students are asked the

right questions (see Gates Foundation, 2010, p. 5). Indeed, the evidence shows that when students report positive classroom experiences, these classrooms tend to achieve greater learning gains, and other classrooms taught by the same teacher do so as well (Gates Foundation, 2010). And student surveys tend to be consistent and reliable across different classes and school years. Student surveys can also provide a rich source of descriptive feedback that teachers can use to aid in their own professional growth. Using multiple sources of information in this way will likely help to paint a more complete picture of teacher effectiveness.

Multiple Ratings of Effectiveness

Many administrators remarked that there should be more than two rating categories so as to distinguish excellent teaching from that which is merely satisfactory. As one administrator put it, “some teachers are well beyond satisfactory and some are just satisfactory, this needs to be acknowledged.” Indeed in the original incarnation of the TPA there were four categories: exemplary, good, satisfactory, or unsatisfactory. It is thus recommended that the TPA revert back to a four point rating scale. One option is to use the original four categories. Another option has been suggested by The New Teacher Project (2010) that centers on the concept of teacher effectiveness: highly effective, effective, needs improvement, or ineffective. This number of categories is large enough to provide teachers with a clear description of their current performance while being small enough to allow for clear distinctions between each level. This will also allow for meaningful differentiation of teacher performance within schools and districts.

More Frequent Evaluations

Many administrators stated that evaluating teachers only once every five years is too infrequent. As one administrator put it, “five years is too long a period in any field for professionals to take a hard look at their practice.” Regardless of their ability level or years of experience, teachers deserve ongoing feedback on their performance. As a teacher’s effectiveness and developmental needs may change over time, more regular evaluations are needed to help satisfy both the competency and professional growth objectives. It is thus recommended that teachers be evaluated every two years. This will help to ensure that teachers receive timely help with their challenges and regular recognition of their successes.

More Regular Follow Up

In terms of professional growth, many administrators commented that the TPA process does not lead to improvements in teacher practice because there is no follow up for teachers who receive a satisfactory rating. As one administrator put it,

once teachers have their copy of the evaluation it is totally within their control whether want to pursue the recommendations or not, unless the TPA is unsatisfactory. Principals/vice-principals cannot mandate additional training/workshops in areas of need therefore the process can be very ineffective.

Therefore, it is recommended that there be more regular follow up with all teachers, regardless of current performance, to ensure that they are taking steps to improve their teaching practice based on the recommendations of their last TPA. Having this regular follow up will ensure that teachers are taking the steps to improve their practice based on the recommendation that administrators spend hours each year carefully crafting. Having TPAs take place every two years

instead of five years will assist in this regard. At this more frequent interval it will be easy to revisit a teacher's last TPA and discuss how those recommendations are being acted upon.

Make Evaluations Significant

Another reason why many administrators felt that the TPA process does not lead to improvements in teacher practice is that it is not used for future personnel decisions. As one administrator put it,

I think this speaks to the fact that as long as you are given a satisfactory appraisal, teachers are satisfied. They understand that it matters little. TPAs are never even mentioned when a candidate is applying for a new job with a new school and administrators are looking for a reference. I have never seen a question that asks "How was their last TPA?"

Therefore it is recommended that TPAs be explicitly factored into important personnel decisions such as hiring, promotion, and possibly compensation. This will communicate the importance of the process to all stakeholders and will ensure that it receives the attention it deserves.

Challenges

A major challenge, especially regarding the recommendations to have unannounced class visits, multiple ratings, use more sources of information, and tie performance on the TPA to hiring, promotion, and compensation is that it may meet stiff resistance from teacher unions. Many administrators commented that there were more disagreements from teachers regarding their evaluations when the four point rating scale was in place and that union involvement often makes the TPA process more difficult. However this need not be the case. In 2009–2010, New Haven Public Schools worked with the New Haven Federation of Teachers to design a new teacher evaluation system that contains many of these recommended elements (New Haven

Public Schools, 2010). To help ensure the system is being implemented fairly, impartial observers review administrators assessments of teachers. This helps to alleviate the anxiety among teachers associated with any new system of evaluation. Hence improvement to teacher evaluation systems can occur where there is meaningful collaboration between school districts and teacher unions.

Conclusion

If teacher evaluation systems such as Ontario's Teacher Performance Appraisal process are to meet their stated objectives of assuring competency and aiding in professional growth, it appears much reform is needed. In its current form, evaluation systems like this are viewed by both teachers and administrators as time consuming but not particularly useful. This is quite a poor public policy combination. It also then begs the question as to why school systems continue such practices. Perhaps one reason is so that political and education leaders can project the appearance of accountability to the public. While this may serve a public relations purpose it imposes a large burden on both teachers and administrators, while not addressing the potential that a more effective system of evaluation can have in aiding professional growth and actually improving teacher practice. An effective evaluation system allows teachers to receive the feedback, support, and recognition they deserve and provides administrators with the information they need to make informed decisions regarding personnel and professional development. By reforming our teacher evaluation systems we can better ensure that all students have access to the most effective teachers possible.

References

- Barnett, J. (2006). *Considering growth in teaching: Teaching perspectives on the Ontario performance appraisal*. Unpublished doctoral dissertation, University of Toronto.
- Barr, A., & Burton, W. (1926). *The supervision of instruction*. New York: Appleton.
- Beerens, D. (2000). *Evaluating teachers for professional growth: Creating a culture of motivating and learning*. Thousand Oaks, CA: Sage.
- Black, C. (2003, April). *Action research and teacher performance appraisal*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Bolger, D., & Vail, H. (2003, May). *Perspectives on "Supporting Teaching Excellence": The Ontario teacher performance appraisal system*. Paper presented at the annual meeting of the Canadian Society for the Study of Education, Halifax.
- Bracey, G. (2004). Value-added assessment findings: Poor kids get poor teachers. *Phi Delta Kappan*, 86, 331–333.
- British Columbia Teachers' Federation. (2012). *2012–2013 Members Guide*. Retrieved from <http://www.bctf.ca/uploadedFiles/public/AboutUs/MembersGuide/guide.pdf>
- Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26, 673–682.
- Coleman, J., Cambell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare.
- Cousins, J. (1995). Using collaborative performance appraisal to enhance teacher's professional growth: A review and test of what we know. *Journal of Personnel Evaluation in Education*, 9, 192–222.
- Creswell, J., & Plano Clark, V. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Creswell, J. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research (3rd ed.)*. Columbus, OH: Pearson.
- Danylewycz, M., & Prentice, A. (1986). Teacher's work: Changing patterns and perceptions in the emerging school system of nineteenth and early twentieth-century central Canada. *Labour*, 17, 57–80.
- Ecker, J. (2001, October 15). Statements by the ministry: Education reform. In Legislative Assembly of Ontario, *Official Report of Debates (Hansard)*, 37th Parliament, 2nd

session. Retrieved from <http://hansardindex.ontla.on.ca/hansardECAT/37-2/L050A-4.htm>

- Ferguson, R., & Ladd, F. (1996). Additional evidence on how and why money matters: A production function analysis of Alabama schools, in H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education*. Washington, DC: Brookings Institution Press.
- Gates Foundation. (2010). *Learning about teaching: Initial findings from the measures of effective teaching project*. Seattle: Bill and Melinda Gates Foundation.
- Greene, J., & Caracelli, V. (Eds.). (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*. San Francisco: Jossey-Bass.
- Goodman, A. (2001). *The evaluation of evaluation: A critical perspective on management aspects of teacher evaluation in Ontario's school systems*. Unpublished doctoral dissertation, University of Toronto.
- Hattie J. (2002). The relation between research productivity and teaching effectiveness: Complementary, antagonistic, or independent constructs? *Journal of Higher Education*, 73, 603–641.
- Haycock, K. (1998). Good teaching matters . . . a lot. *Thinking K–16*, 3, 3–14.
- Houston, S.E., & Prentice, A. (1988). *Schooling and scholars in nineteenth-century Ontario*. Toronto: University of Toronto Press.
- Johnson, S. (1990). *Teachers at work: Achieving success in our schools*. New York: Basic Books.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Newbury Park, CA: Sage
- Joyce, B., Calhoun, E., & Hopkins, D. (1999). *The new structure of school improvement: Inquiring schools and achieving students*. Buckingham, UK: Open University Press.
- Kauchak, D., Peterson, K., & Driscoll, A. (1985). An interview study of teachers' attitudes towards teacher evaluation practices. *Journal of Research and Development in Education*, 19, 32–37.
- Larsen, M. (2009). Stressful, hectic, daunting: A critical policy study of the Ontario Teacher Performance Appraisal system. *Canadian Journal of Educational Administration*, 95, 1–44.

- Lawton, S., Hickox, E., Leithwood, K., & Musella, D. (1988). Performance appraisal in Ontario Schools. In E. Hickox, S. Lawton, K. Leithwood, & D. Musella (Eds.), *Making a difference through performance appraisal* (pp. 13–41). Toronto: The Ontario Institute for Studies in Education Press.
- Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Marshall, K. (2009). *Rethinking teacher supervision and evaluation: How to work smart, build collaboration, and close the achievement gap*. San Francisco: Jossey-Bass.
- Medley, D., & Coker, H. (1987). The accuracy of principals' judgements of teacher performance. *Journal of Educational Research*, 80, 242–247.
- Middlewood, D. (2002). Appraisal and performance management. In T. Bush & L. Bell (Eds.), *The principles and practice of Educational management* (pp. 119–134). London: Paul Chapman.
- Miles, M., & Huberman, A. (1994). *Qualitative data analysis: A sourcebook for new methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Miller, T. (2009). *Investigating elementary teachers' perceptions about and experiences with Ontario's Teacher Performance Appraisal system*. Unpublished doctoral dissertation, University of Toronto.
- Ministry of Education. (2002). *Supporting teaching excellence: Teacher performance appraisal manual and approved forms and guidelines*. Toronto: Queen's Printer for Ontario.
- Ministry of Education. (2010). *Teacher performance appraisal: Technical requirements manual*. Toronto: Queen's Printer for Ontario.
- Morris, L. (1930). *The single salary schedule: An analysis and evaluation*. New York: Teachers College, Columbia University.
- New Haven Public Schools. (2010). *Teacher evaluation and development: Introduction to the process*. New Haven, CT: New Haven Public Schools.
- Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237–257.
- Odden, A. (2011). *Strategic management of human capital in education*. New York: Routledge.
- Osborne, B. (1987). *Appraisal for professional growth*. (EDRS Document Reproduction Service No. ED294319)
- Peterson, C. (1982). *A century's growth in teacher evaluation in the United States*. New York: Vantage.

- Peterson, K. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices* (2nd ed.). Thousand Oaks, CA: Corwin
- Porter, A., Youngs, P., & Odden, A. (2001). Advances in teachers assessments and their uses. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 259–297). Washington, DC: AERA.
- Prentice, A., & Theobald, M. (1991). Teacher's work: Changing patterns and perceptions in the emerging school systems of nineteenth and early twentieth-century central Canada. In *Women who taught: Perspective on the history of women and teaching*, (pp. 136– 159). Toronto: University of Toronto Press.
- Rice, J. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.
- Rivkin, S., Hanushek, E., & Kain, J. (2001). *Teachers, schools, and academic achievement* (Working Paper 6691). Cambridge, MA: National Bureau of Economic Research.
- Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 471–458.
- Rowe, B. (2000, April). *The influence of teacher efficacy and readiness for self-directed learning on the implementation of a growth-oriented teacher performance appraisal process*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.
- Salisbury, A. (1912). *School management: A textbook for normal schools*. Toronto: Educational Book Co.
- Sergiovanni, T. (1992). *Moral leadership: Getting to the heart of school improvement*. San Francisco: Jossey-Bass.
- The New Teacher Project. (2010). *Teacher evaluation 2.0*. New York: author.
- Thomas, E. & Wingert, P. (2010, March 15). Why we must fire bad teachers. *Newsweek*, 155(11), 24–28.
- Vold, D. (1985). The roots of teacher testing in America. *Educational Measurement: Issues and Practice*, 4, 5–7.

- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York: The New Teacher Project.
- Whitehurst, G. (2002, March). *Scientifically based research on teacher quality: Research on teacher preparation and professional development*. White House Conference on Preparing Tomorrow's Teachers.
- Wise, A., Darling-Hammond, L., McLaughlin, M., & Bernstein, H. (1984). *Teacher evaluation: A study of effective practices*. Santa Monica, CA: RAND.
- Wolf, R. (1973). How teachers feel toward evaluation. In E. House (Ed.), *School evaluation: The politics and process* (pp. 156–168). Berkeley, CA: McCuthan.

Appendix: Full Survey Results

This section presents each survey question and its results. In addition to the quantitative results, representative comments are provided where appropriate.

1) How much training have you received in conducting Teacher Performance Appraisals (TPAs)?

Forty-one percent of participants said they received a moderate amount of training, 34% said they received a great deal or a lot of training, while 25% said they received little or no training at all. Comments indicated that training appears to largely consist of workshops by school boards and the OPC that are entirely voluntary. This may help explain the wide range of responses.

	Response Percent	Response Count
A great deal	12%	21
A lot	22%	39
A moderate amount	41%	72
A little	22%	38
None at all	3%	6

2) Please describe the training you have received in conducting TPAs.

Participants indicated that they received mandatory training when the Ministry of Education first introduced TPAs to the province's schools but, since then, the only training that have been provided are voluntary workshops put on by school boards and the OPC. In terms of the content of the training, administrators that do attend the voluntary workshops are being

trained in the technical aspects of the TPA, but not necessarily in how to accurately assess and evaluate teachers. As one participant put it, "I was trained in what all of the aspects of the TPA are. There was not so much information about how to conduct the meetings or observations."

3) How useful was the training you received in conducting TPAs?

Thirty-six percent of participants said they found the training they received extremely or very useful, 38% found it moderately useful, while 26% found it either slightly or not at all useful.

	Response Percent	Response Count
Extremely useful	9%	15
Very useful	27%	47
Moderately useful	38%	65
Slightly useful	21%	35
None at all useful	5%	9

The comments to this question reflected the wide range of responses. In the answers to both this and the previous question, participants seemed to indicate that there was a lack of training in dealing with unsatisfactory TPAs. As one participant put it, "no training in unsatisfactory TPAs. Superintendents seem to be trying to say something they don't want to directly say in our training." It is thus implied that a possible reason many administrators are not receiving training in dealing with unsatisfactory TPAs is that they are being discouraged from giving an unsatisfactory rating to teachers.

4) What other preparation do you undertake prior to each performance appraisal?

The most common additional preparation administrators reported undertaking was meeting with the teachers they were going to evaluate. This involved reviewing the process and requirements of the TPA process with the teacher and establishing timelines for classroom observation and conferences. The second most common response of administrators was that they reviewed the TPA manual and any other policy documents related to the requirements of the evaluation process. Although not one of the most common responses, some reported that they did not engage in any additional preparation. The exact reasons for this are unclear, but one administrator gave the response “none. This job is too busy and the TPA’s are just an added workload issue that are not really used for anything.” This may indicate that some administrators do not view conducting TPAs as an important part of their job when compared to their many other responsibilities.

5) How long do you typically spend observing a teacher during a classroom observation?

Administrators reported spending an average (mean) of 86 minutes observing teachers during the classroom observation portion of the TPA. Answers ranged from a low of 20 minutes to a high of 300 minutes.

6) How adequate are the classroom observations in assessing teacher practice?

The majority of participants (51%) responded that the classroom observations were moderately adequate in assessing teacher practice. Twenty-six percent reported that the classroom observations were extremely or very adequate, whereas 23% found them to be slightly or not at all adequate.

	Response Percent	Response Count
Extremely adequate	5%	9
Very adequate	21%	37
Moderately adequate	51%	89
Slightly adequate	20%	34
Not at all adequate	3%	5

Many administrators felt that the pre-planned nature of the classroom observation did not allow for an accurate or thorough assessment of a teacher's regular teaching practice. Here are some representative comments:

The "formal" classroom observation is the least helpful of the sources of data for a performance appraisal. In my experience, marginal teachers are able to put together a solid lesson that does not reflect their daily teaching practice.

Observing a single class does not give a very thorough picture of a teacher's practice. A teacher can always put on a good show.

7) Evaluate this statement: "More classroom observations would allow for a more accurate and useful teacher appraisal."

The majority of participants agreed that more classroom observations would allow for a more accurate and useful teacher appraisal, with 64% stating they strongly agree or agree, 24% saying they disagree or strongly disagree, and 12% undecided.

	Response Percent	Response Count
Strongly agree	33%	58
Agree	31%	54
Undecided	12%	20
Disagree	20%	35
Strongly disagree	4%	7

While the majority of administrators agreed that more classroom observations would allow for a more accurate and useful teacher appraisal, many felt that there was simply no more time in their busy schedules to allow for this. As one administrator put it, “yes, however where is this time going to come from on the part of the administrator?”

Some administrators also felt that more observations would not be useful unless they were different than the current scheduled visits. They felt that the observations needed to be unscheduled in order to provide an accurate picture of the teacher’s daily teaching practice. Here are some representative comments:

More is not really the issue. This would only be useful if we get to pick and choose where and when and be permitted to have some observations be random.

Especially drop in visits that aren’t scheduled, so the practice you are observing is reflective and true to the daily practice.

8) Besides classroom observations, what other information do you consider when making your assessment?

Administrators reported using a variety of sources other than classroom observations when making their assessment including the teacher’s lesson plans, unit plans, assessments,

notebooks, conduct during staff meetings and professional development sessions, relations with staff and students, and what was often termed “contribution to the school community.”

Interestingly, many administrators reported utilizing sources of information outside of those strictly dictated by the current TPA process. This included informal visits to the teacher’s classes throughout the year, report card data including the teacher’s pass-fail rates, and anecdotal input from curriculum leaders or department heads. It appears that administrators may be using these “outside” sources of information because they feel that the current sources of information available to them within the guidelines of the TPA process (e.g. scheduled classroom observations) are inadequate.

9) How often do you conduct performance appraisals for teachers that teach subjects outside your own area of expertise?

Forty-five percent of administrators reported that they conduct performance appraisals for teachers that teach subjects outside their area of expertise always or most of the time, 17% reported conducting them about half the time, and 38% reported conducting them once in a while or never.

	Response Percent	Response Count
Always	19%	33
Most of the time	26%	46
About half the time	17%	30
Once in a while	29%	50
Never	9%	15

Many administrators took umbrage at the question, seeming to indicate that it challenged their ability to be an effective evaluator. Many commented that subject area expertise is not required in order to effectively evaluate a teacher. Here are some representative comments:

Effective instruction is my area of expertise. If you are suggesting by this question that assessors should have taught the subject of the teacher engaged in the TPA, you are asking a completely irrelevant question.

It is my belief that teachers teach. What they teach is not the skill set so I don't think I'm ever evaluating outside my area of expertise because teaching is my expertise.

10) How confident are you in conducting performance appraisals for teachers that teach subjects outside your own area of expertise?

Seventy-one percent of administrators reported being extremely or very confident conducting performance appraisals for teachers that teach subjects outside their own area of expertise, 25% reporting being moderately confident, 4% reporting being slightly confident, and no administrators reporting being not at all confident.

	Response Percent	Response Count
Extremely confident	23%	40
Very confident	48%	84
Moderately confident	25%	43
Slightly confident	4%	7
Not at all confident	0%	0

In the comments to this question many administrators reiterated their belief that subject area expertise is not required in order to effectively evaluate a teacher.

11) How confident are you in conducting performance appraisals for teachers that teach subjects that are in your own area of expertise?

Ninety-seven percent of administrators reported being extremely or very confident conducting performance appraisals for teachers that teach subjects that are in their own area of expertise, 3% reported being moderately confident, and no administrators reported being slightly or not at all confident.

	Response Percent	Response Count
Extremely confident	63%	107
Very confident	34%	58
Moderately confident	3%	5
Slightly confident	0%	0
Not at all confident	0%	0

When we compare the responses to this question with those in the previous question it appears that despite the comments about subject area expertise being unnecessary, administrators feel much more confident when evaluating teachers that teach subjects that are in their own area of expertise.

12) How soon after the classroom observation does the post-observation meeting usually take place?

The average (mean) response to this question was 4.33 days, which seems to meet the ministry requirement that the meeting be held “as soon as possible after the classroom observation.” Responses ranged from a low of 1 day to a high of 30 days.

13) How often have teachers disagreed with your assessment?

Almost all administrators reported rarely having to deal with teachers disagreeing with their assessment, with 97% reporting that this happened only once in a while or never, 1% (n=2) reporting it happened about half the time, and 2% (n=3) reporting it happening always or most of the time.

	Response Percent	Response Count
Always	1%	2
Most of the time	1%	1
About half the time	1%	2
Once in a while	43%	72
Never	54%	92

Many administrators commented that only time they had teachers disagree with their assessment was when they were given an unsatisfactory rating, which does not appear to happen very often. Administrators also reported that there were more disagreements when the previous four point rating scale of unsatisfactory, satisfactory, good, and exemplary was in place.

It also appears that one possible reason that there was so little disagreement was that once a teacher achieves the satisfactory rating, the TPA is not used for future personnel decisions. As one administrator put it,

I think this speaks to the fact that as long as you are given a satisfactory appraisal, teachers are satisfied. They understand that it matters little. TPAs are never even mentioned when a candidate is applying for a new job with a new school and administrators are looking for a reference. I have never seen a question that asks "How was their last TPA?"

14) When teachers have disagreed with your assessment, how have you handled this?

When they have faced teachers who disagreed with their assessment, administrators almost universally reported engaging with these teachers in a dialogue and presenting the evidence for their assessment. Teachers were then given the opportunity to respond, sometimes changes to the evaluation were made and in some cases they were able to come to a mutual understanding. However despite this approach, many administrators reported that the superintendent and union officials needed to be called in to deal with the issue. Some stated that they received pressure from union officials, which made their job as an evaluator much more difficult. As one administrator put it, “the federation is a daunting presence in such circumstances. It takes a lot of time and the time is worth it because it is so important. But it is very stressful and the federation can make your life just plain miserable.”

15) Evaluate the following statement: “My role as an evaluator has negatively affected relations with my teaching staff.”

Most administrators did not feel that their role as an evaluator negatively affected relations with their teaching staff, with 83% stating that they disagree or strongly disagree, 12% undecided, 4% agree, and no administrators stating that they strongly agree.

	Response Percent	Response Count
Strongly agree	0%	0
Agree	4%	7
Undecided	12%	21
Disagree	48%	82
Strongly disagree	35%	60

Most administrators stated that they embraced the evaluation role as part of their job and that in fact it has often enhanced their relations with their teaching staff. Many viewed it as an opportunity to provide coaching and mentorship and to build better relationships with their teachers. It was also used as an opportunity to learn more about and celebrate the practice of the excellent teachers in their schools. The only instances where it appeared to harm relations were with those teachers that received an unsatisfactory rating. Here are some representative comments:

In most cases it provided me opportunities to celebrate teachers' successes which has enhanced the relationships.

My relationship with the teacher who received an unsatisfactory appraisal was severely impacted. However, other teachers appreciate that someone is holding the staff accountable.

16) Evaluate the following statement: "The TPA process accurately assesses the practice of teachers."

Administrators are divided on the question as to whether the TPA process accurately assesses the practice of teachers, with 41% stating they strongly agree or agree, 28% undecided, and 32% stating they disagree or strongly disagree.

	Response Percent	Response Count
Strongly agree	4%	6
Agree	37%	62
Undecided	28%	46
Disagree	25%	41
Strongly disagree	7%	12

Accordingly, administrators gave a range of responses in their comments. Many felt that the TPA process is only as good as the administrator conducting it. Others noted that the process is good at highlighting the work of great teachers but not at addressing mediocre ones, partly as a result of union involvement. Still, others felt the process was too contrived and artificial to be of use.

17) How much has teacher practice improved in your school as a result of the TPA process?

Overall, most administrators felt that there has not been substantial improvement of teacher practice in their schools as a result of the TPA process. This was reflected in both the answer to the question and the comments that were provided. Just 9% indicated that teacher practice had improved a substantial amount, 34% a modest amount, 36% a small amount, and 20% not at all.

	Response Percent	Response Count
A substantial amount	9%	15
A modest amount	34%	55
A small amount	36%	59
Not at all	20%	33

When we compare these results with those found by Lawton, Hickcox, Leithwood, and Musella in 1988 when they asked this question to teachers and superintendents in Ontario, we see that administrators are in much closer agreement with teachers who felt that the teacher evaluation system at the time did not result in substantial improvement in teacher practice or performance.

Most administrators appeared not to view the TPA process as a major vehicle for improving teacher practice. Instead, many indicated that improvements in teacher practice were the result of ongoing professional development in the school during PLC time, PD days, staff meetings, and coaching and mentoring amongst administrators and staff. Many administrators commented that because there is no follow up for teachers who get a satisfactory rating, improvements in practice usually only occur with those that receive an unsatisfactory rating.

Here are some representative comments:

It is not the TPA that has improved practice, it has been the PLCs, Critical Learning Pathways, staff discussions and meetings, and other such PD that has made an impact.

Once teachers have their copy of the evaluation it is totally within their control whether want to pursue the recommendations or not, unless the TPA is unsatisfactory.

Principals/vice-principals cannot mandate additional training/workshops in areas of need therefore the process can be very ineffective.

18) On average, how many hours per school year do you spend on conducting TPAs and its related activities?

Administrators reported spending an average (mean) of 58 hours per year on conducting TPAs and its related activities. Thus it appears that TPAs consume a substantial portion of administrative time. Answers ranged from a low of 10 to a high of 200.

19) Evaluate the following statement: "Conducting TPAs is an effective use of my time."

Administrators were more likely to say that conducting TPAs is an effective use of their time, with 47% stating that they strongly agree or agree, 21% undecided, and 32% disagree or strongly disagree.

	Response Percent	Response Count
Strongly agree	8%	14
Agree	39%	65
Undecided	21%	35
Disagree	22%	36
Strongly disagree	10%	17

Many administrators commented that while they enjoyed observing teachers and having an opportunity to engage them about their teaching practice, the time spent on paperwork and other bureaucratic aspects of the evaluation was not as worthwhile. Others felt that without more than two rating categories or follow up for those who receive satisfactory ratings the process was not as effective as it could be. Yet many still viewed the process as valuable in that it appears to be one of the only vehicles to address poor teaching.

20) How do you feel the TPA process could be improved?

By far the most common suggestion from administrators was to have more classroom observations. Many administrators felt that basing the TPA on a single classroom observation was insufficient. In addition, administrators stated that, to get a true representation of a teacher's teaching practice, at least some of the observations should be unannounced. Administrators also stated that evaluating teachers only once every five years was too infrequent. Many also remarked that there should be more than two rating categories so as to distinguish really excellent teaching from that which is merely satisfactory. Another common suggestion was to streamline the number of competencies that teachers must be evaluated on as many are very

similar. In addition, some administrators also felt that there should be a cap on the number of TPAs they must perform in a given year so that they can give enough attention to each one. Here are some representative comments:

More classroom observations both invited by the teacher under performance appraisal as well as impromptu observations as a principal.

I would like to see one observation planned and the other observation unplanned. This would reflect what is actually going on in the classroom.

Five years is too long a period in any field for professionals to take a hard look at their practice. I would suggest having TPAs every 3 years.

Change satisfactory and unsatisfactory to a tiered ranking. Some teachers are well beyond satisfactory and some are just satisfactory, this needs to be acknowledged.

21) Please describe your gender.

Sixty-nine percent of participants identified themselves as female and 31% as male.

	Response Percent	Response Count
Female	69%	113
Male	31%	51

22) In what school level(s) have you been an administrator?

Sixty-one percent of participants stated that they have been administrators at elementary schools, 29% at secondary schools, and 10% stated that they have been administrators at both school levels.

	Response Percent	Response Count
Elementary	61%	101
Secondary	29%	48
Both	10%	17

23) What is your current administrative position?

Sixty-four percent of participants stated that they are currently a principal and 36% indicated that they are currently a vice-principal.

	Response Percent	Response Count
Principal	64%	105
Vice-principal	36%	60

24) Describe the geographic setting of your school.

Forty-four percent of administrators identified their school as urban, 36% as suburban, and 20% as rural.

	Response Percent	Response Count
Urban	44%	72
Suburban	36%	59
Rural	20%	34

25) How many years have you been an administrator?

Participants indicated that they had an average of 8.35 years of administrative experience. Answers ranged from a low of 1 year to a high of 24 years.

26) Prior to becoming an administrator, for how many years were you a teacher?

Participants indicated that prior to becoming administrators they had been teachers for an average of 13.86 years. Answers ranged from a low of 5 years to a high of 30 years.