From Watchful Eye to Bureaucratic Formality:
The Evaluation of Teacher Performance
in Ontario from 1870 to 1999

by

Jennifer Gail Watt

A thesis submitted in conformity with the requirements for the degree of Masters of Arts
Department of Theory and Policy Studies
Ontario Institute for Studies in Education of the University of Toronto

© Copyright by Jennifer Gail Watt 1999
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Abstract

From Watchful Eye to Bureaucratic Formality: The Evaluation of Teacher Performance in Ontario from 1870 to the present
M.A. 1999
Jennifer Watt
Department of Theory and Policy Studies
Ontario Institute of Studies in Education of the University of Toronto

This document presents a synthesis of trends from 1870 to 1999 in teacher evaluation in the Ontario school system which includes four eras which overlap each other and are not so fixed as these rough time frames might suggest: the “inspection era” (1870s-1900s); the “scientific evaluation” era (1920s-1930s); the “humanist assessment” era (1960s-early 1970s); and the “business oriented, back to basics” performance review era of the 1990s. It is argued that explicit and implicit purposes of teacher evaluation have not significantly changed and that classroom observation is still the main source of data in evaluating teacher effectiveness.

An examination follows of 20 teacher reports dating from 1982-1998, from one Ontario school board. Issues in the data include: the effect of gender in the report, the lack of consistent use of criteria, the non-critical tone of the reports and the use of questionable criteria.

The author concludes with a review of new research advocating increased teacher autonomy along with a critique of the conservative political climate that resists progressive changes.
Setting the Stage-Problem Statement

In my first year of teaching, I remember that moment of dread when I was informed, by memo, of my impending teacher's evaluation. How was I to demonstrate my dedication, professionalism and talent in one 75-minute lesson observed by my principal? I was given a prescribed format in which to summarize my lesson for the administrator. Headings on the required form listed topic, resources, strategies, modifications and assessment and evaluation as the basis of the 15 minute pre-conference that was to occur before the actual event. I recall an inordinate amount of effort being placed into the preparation for this one lesson. Admittedly with insufficient regard for professionalism, other teachers coached me on their technique of misleading their students into believing that the principal's purpose in attending the class was to monitor student performance. These teachers assured me that this was a foolproof way to ensure the student's best behaviour resulting in a glowing report from the principal.

Teachers enjoy relating, and perhaps exaggerating, both the comedic and horrific stories of ill-fated evaluations that frequently involved students who declared in the middle of a flawless lesson, "Why don't you always teach us this way?". The stories varied from the humorous to the pathetic involving rigid inspectors and principals who commented, in an official written report, about the perspiration stains on the teacher's shirts, the lack of recent student work displayed on the bulletin boards or the teacher's inability to use coloured chalk effectively. Other colleagues claimed that they taught the same perfected generic lesson to new administrators year after year. Presumably, this masterpiece of a lesson plan never failed to impress.

Depending on a teacher's prior experiences, he or she will have a range of emotions regarding classroom observation. At one extreme is resentful acceptance while the other extreme is an enthusiastic willingness to display one's professional expertise and skills. I can honestly state that I can't recall in detail anything about my evaluations over the past ten years except the stress that accompanied them. And I was in good company - teachers who had been in the profession over 20 years seemed unduly anxious by this procedure. I felt somewhat like a primate at the zoo, with a presumed expert in animal behaviour checking off boxes and taking copious notes regarding my conduct. In other cases, the evaluator just "dropped by", sat for a few minutes to observe or converse with the students about the lesson and then they left. No matter how glowing and
complimentary the report, I always felt a bit uneasy and suspicious of this process. Educational studies have consistently reported that summative evaluations including classroom observations and written reports represent a psychological endpoint or final decision that causes a sense of anxiety and seriousness, whether or not these feelings are justified (Daresh & Playko, 1995). How exactly did these "endpoint" procedures benefit my students or aid my professional growth? It is with this set of concerns that I decided to research the historical development of teacher evaluation in Ontario, the major trends found in its purposes and techniques, and to what degree, if at all, teacher evaluation serves as a valuable tool of education.

Research Methodology

My main purpose in reviewing both the historic literature as well as critiquing modern documents from one specific urban public school board in eastern Ontario is to assess teacher evaluation from both historical and contemporary perspectives. A comparison of the main educational paradigms in Ontario from the 1870s to specific information about the trends and experiences of teacher evaluation found in each time period will reveal any viable connection between the historical and contemporary situations. Does a period of centralization, consolidation and imposed socialization reflect itself in a specific policy such as teacher assessment? When pedagogical attitudes change and education becomes a more humanistic and progressive endeavor, do administrators radically alter teacher assessment policies to manifest these changes?

Archival regulations and minutes from meetings at the Toronto School Board, annual inspector's reports to Directors of Education and readings on the history of education in Canada and Ontario were used to gain an understanding of major trends. Texts on supervision and teacher assessment were also considered in each time period: Appleton, 1897; Salisbury, 1912; Barr, 1938; Brown, 1948; Flower, 1958; Wilson, 1966; Netzer, 1970; Hoy & Miskel, 1996. My concern was to try to get a general picture of the changes in teacher assessment in a period of over a century before undertaking an analysis of contemporary teacher evaluation reports. When information was lacking on Canadian trends, American sources are used. I have ignored some details and interesting comparative questions and perhaps blurred exact differences between teacher assessment in the United States and Canada. My main objective was to examine educational
attitudes and how they correspond to policies and practices that were found in nearly all public schools in North America, both elementary and secondary, in an attempt to detect the common tendencies that have seemed to act upon teacher evaluation. A discussion of findings will attempt to reveal how the stated policy of teacher assessment of each era mirrors educational paradigms.

From the sources used, I identify four distinct patterns that have characterized educational theorists' attitudes towards the assessment of teacher performance from the 1870s to the present and argue for the cessation of policies and procedures of teacher assessment that are largely questionable in terms of validity, reliability and effective educational practice. My evidence concerning contemporary policies used in teacher assessment is derived from one specific board in Ontario. By analyzing 20 written teaching reports resulting from the most current policy documents on teacher evaluation in one specific board and through several informal interviews with teachers, department heads, principals and supervisory officers employed with that Board, I attempt to answer the following questions: "What are the stated purposes of teacher evaluation?", "How have purpose and policies changed?", "Are the purposes and policies reflected in the product of teacher assessment reports?" and "Have the procedures that have been used in the past fundamentally changed?" Lastly, and most importantly, "Do changes in purposes, policies and procedures reflect the major paradigms in educational and societal thought in Ontario?". Because the contemporary reports have been selected at random, and the nature of the written report as an official and confidential document, there are bound to be distortions in what was recorded. However, the written teacher evaluations serve as useful data concerning overall trends and highlight specific concerns in the reporting of a classroom observation as outlined by Board-generated guidelines and policies. Some anecdotal observations will be made in regards to teacher assessment procedures based on my personal experience and informal discussions as a teacher and department head within the same Board.

Conceptual Framework

The literature on Ontario educational history is vast, offering a multitude of conceptual frameworks to consider. The literature on historical trends in teacher assessment is lacking, perhaps due to the fact that teacher files have not been maintained in archives until the 1950s in Toronto and
surrounding boards. Secondary sources remain the main source of information. Several frameworks that I found useful in undertaking an examination of how teacher assessment is linked to overall educational movements are provided by educational and political theorists Beach and Reinhartz (1989), Candoli (1995) and Manzer (1994). All of these researchers have developed conceptual settings that attempt to define the evolution of the educational system as it responded to societal needs, values and attitudes.

Carl Candoli (1995), former Deputy Commissioner of Education for Texas, studied the historical development of the duties of superintendents, concluding that major shifts such as the transition to a industrial society, the principles of scientific management, the human relations approach and the behavioural school of management all significantly changed supervisory roles. Candoli sees these American historic changes as dramatic and tumultuous, not the gradual evolution he initially hypothesized. The way supervisors were expected to monitor teachers altered according to changing educational beliefs. Industrialism and the resulting beliefs in scientific management theoretically turned supervisors into managers who were checking for efficiency in their employees. The human relations and behaviouralist approach was meant to transform the supervisor into a partner and coach of the teacher through the assessment process. Candoli’s analysis shows that developing educational philosophies led to new expectations in supervision. I will argue that the “dramatic and tumultuous” change in the United States was more moderate in tone and slower in pace in the Canadian educational experience due, in part, to what Ronald Manzer identifies as our nation’s inherent liberalism.

Beach and Reinhartz’ (1989) study also reveals that schools are a reflection of the values and attitudes of each generation. They align these major historical trends to educational supervision. Beach and Reinhartz label the historic phases in American education as: the colonial period (1600-1865); the state and national period (1865-1910); the scientific and organizational period (1910-1920); the professional and bureaucratic period (1920-1935); the progressive and cooperative period (1935-1955); the change oriented period (1955-1970); the clinical and accountability period (1970-1980); and the managerial period (1980 to the present). The ideas of these eras move in and out of each other, inform and illuminate each other. In short, there are no rigid boundaries between them. These fluid historic phases, while more focused on specific legal changes in educational
supervision, are more or less mirrored in the Canadian experience. Canada struggled to create a public education system that focused on fostering civic and national identity. The desire for legitimizing a new and expensive system paired with changing technologies and a more diversified economy led to the scientific and organizational periods in both Canada and the United States. Increased progressivism in education stirred significant transformations in many schools in North America. Accountability and management are terms that are also gaining increasing acceptance in the Canadian educational lexicon. I have chosen to condense four of Beach and Reinhartz' periods (the scientific and professional and the clinical and managerial) into two single entities for the purpose of stressing my assertion that practices in Ontario educational supervision tend to combine the attributes of these phases for longer periods of time.

Canadian political scientist Ronald Manzer's (1994) conceptualizes how the education system responded to political change through the formation of specific policies. Manzer focuses on the school system as a major arena in the struggle for political power. It is within the educational sphere that citizens of a political democracy strive for "civic virtue, economic wealth, social integration, and cultural survival" (p. 3). Manzer sees Canadian public schools as based on liberal tenants that have changed their focus in four major eras. These eras are: political liberalism, from the foundation of the public school system to the mid 19th century; economic liberalism and accessibility in the early 20th century; ethical liberalism in the 1960s; and technological liberalism and the quest for educational excellent in the 1990s. Manzer's analysis is on legal policies in all areas of education and does not consider supervision and teacher assessment. The Canadian tendency towards liberalism and the philosophy of middle ground compromise in aspects of politics, law and culture has been noted by many Canadian theorists (e.g. Saul, 1998) and political scientists, such as Manzer, as an important part of our national identity. This fundamental liberalism may distinguish our educational experiences from the United States and helps to explain our tendency to move more gradually in accepting and implementing new educational ideas. Perhaps, our liberalist educational views could even be one of the main explanations for the lack of formal testing of non-tenured and tenured teachers.

In reviewing the historical sources, I was struck at how contentious the issue of the formal evaluation of teachers has been since the days of the dreaded, all-powerful inspector's annual visits.
Changes in societal beliefs have contributed to defining and redefining the purposes, procedures and foci of the evaluation of public school teachers but has done little to eliminate the debate on the usefulness of the procedure. These historic changes include increased centralization, standardization of teacher training and certification, the development of teaching unions and the demand for educational accountability. All these factors have resulted in changes in teacher evaluation policies. Attempts at a more centralized school system with standardized curriculum have had a tendency to be reflected in teacher evaluation policies that, in theory, were punitive and authoritarian in nature. As teachers gained more professional status and strength in union organizations, evaluation policies seemed to have become a more collaborative and supportive endeavor.

In undertaking an analysis of the trends in teacher supervision in Ontario, the three noted historic frameworks are useful in their identification of educational paradigms. Candoli (1995) as well as Beach and Reinhartz (1989) focus on the changes in the American supervisory role while Manzer (1994) critiques Canadian political patterns as they were manifested in educational policy. The historical overview I offer applies specifically to the stated policies, procedures and practices of teacher performance assessment. A combination of the Beach and Reinhartz categorization of educational periods along with Manzer's identification of the Canadian liberalist tradition will be the basis of this analysis. Although my overview is organized chronologically, according to which trends I have identified as seeming the most prominent at the time, one or more of these trends might have been invoked at any time, depending on circumstances. In other words, the situation was always more complex that this, or any, conceptual framework might appear. Nonetheless, the four historical trends may offer a useful heuristic for those attempting to rethink the evolution of teacher evaluation policies and procedures.

The evaluation of teacher performance has been and still is a process that defines proper professional behaviour in a specific societal setting (Varga, 1991). The exact terminology used to define the process of a superior observing and judging a teacher's performance gives important clues to the educational views of a particular era. As the historic examples illustrate, "inspections" were frequently autocratic excesses; "evaluations" implied an unfounded scientific basis for the resulting judgments made of the teacher; "assessments" were a softer, process-oriented and humanist affair;
"reviews" combined both the formative and summative purposes that are commonly found in the arena of business management. “Teacher testing” is now to be introduced in Ontario as yet another shift in supervision and teacher assessment towards qualitative accountability (Toronto Star, April 26, 1999). These changes in the language of teacher evaluation are revealing when reviewing developments in the Ontario public school system beginning, with its foundation, in the 1870s.

Regardless of new trends and attitudes being reflected in new policies, the procedure of formal monitoring of teachers by their superiors has continued in much the same fashion as a century ago. I feel that there may be two main reasons revealed in primary accounts and in secondary analysis, one explicit and one implicit, that have ensured the continuation of formal teacher evaluation reports based solely on classroom observation. The first explicit reason is the pervasive belief held by some administrators, educators and even the general public that teacher evaluation is an important way to better education. Teacher evaluation is thought by some to ensure that effective teachers are given the responsibility of facilitating learning for students. Ineffective teachers should be revealed, remediated, monitored or even dismissed since they deter from educational goals. This belief seems to continue in some educational discussions despite conflicting research as to the effectiveness of teacher evaluation reports either as a tool to help teachers improve their practice or, in the most extreme cases, in helping to remove incompetent teachers (Johnson, 1978). Secondly, and perhaps more significantly, mandatory evaluation by classroom observation of contract (versus probationary) teachers may have an implicit purpose in maintaining a hierarchy of roles and a power structure (superintendent, principal, department head, teacher) that remains the norm in North American school systems (Wiebe, 1969; Finkelstein, 1974; Haefele, 1993).

These two potential reasons, one pedagogical and one political, have helped to guide teacher evaluation policies in the public schools. If the reasons largely remain the same, what has changed since the 1870s in formal teacher assessment? One could speculate that the criteria for what makes an effective teacher have altered because of changing cultural attitudes. While the criteria for what makes an effective teacher have changed, purposes of teacher evaluation have seemed to stay constant over time while the main procedure used to collect the data has not evolved either. Classroom observation and the subsequent written report first employed over 120 years ago remain as the most important tools used in teacher performance assessment in Ontario (Musella, 1988;
Educational Paradigms and Teacher Evaluation

From a review of the literature on teacher evaluation, I identify four overall educational paradigm shifts in Canadian education. The following dates provide a helpful overall framework of the major changes regarding the monitoring of teachers through formal, written evaluation. To emphasize an earlier point, I recognize that the beliefs of one era often overlapped and flowed into the next, causing a unique hybrid of values. For ease of historical analysis, however, the following four eras will be considered as separate units, complete with their own unique terminology, beliefs, standards and methods for teacher evaluation.

The years from the 1870s through to the early 1920s could be named the "inspection era" in teacher evaluation and coincides with what Ronald Manzer (1994) defines as the "political liberalism" era of education. This was a time of considerable economic growth in North America. The corresponding growth in the education system resulted in fundamental political struggles over schools, teachers and curriculum. The 1920s to the late 1950s reflect society's increasing focus on scientific testing and new economic realities. Canadian education was in its "economic liberalist" stage (Manzer, 1994). During this "evaluation era" of teacher competency, the United States, and Canada to a lesser degree, produced hundreds of psychological studies, statistical means and rating charts to assess student performance and teacher effectiveness (Barr, 1929). The "assessment era", which gathered strength in the late 1950s and carried through to the mid 1970s, focused on humanist and democratic educational theories in what Manzer defined as the "ethical liberalism" stage of Canadian education. Manzer concludes that contemporary "technological liberalism" is a political quest for educational excellence that is closely linked to business interests. The current era is more difficult to define but the focus seems to be a peculiar combination often found in the business realm of top down inspection with humanist terminology softened by humanist terminology. What follows is an elaboration on these general points.
1870s-1920s: The Inspection Era of Supervision

"The inspection era" was a trend in teacher assessment that was formalized in the 1870s with the formation of publicly funded schools in Ontario and continued to dominate many beliefs and practices concerning teacher evaluation into the early decades of the twentieth century. It was a time characterized by a continuing political struggle for a centralized education system. The Reverend Egerton Ryerson and other Ontario reformers used legislative means to get previously autonomous school administrations into a more uniform provincial education system. It was hoped that parents could be assured that teachers were adequately trained and monitored. The shared aspirations and anxieties of a new middle class for an education system that would equal private schools in granting their children an intellectual and economic footing with the elite in society united the movement led by Ryerson to establish universal tax-supported common schools (Phillips, 1957). Ryerson successfully promoted the importance of a public school system in maintaining and improving a civilized nation: "Education is a public good; ignorance is a public evil. What affects the public ought to be binding upon each individual composing it" (Ryerson cited in Houston, 1972). Provincial educational leaders continued to solidify their power largely through legislative means during this growth period, despite protests from local trustees and promoters of private schools (Danylewycz, 1986). By 1871, the Ontario legislature had passed an education act that required boards to provide schooling through local taxation. One critical measure used to maintain power over the newly accountable teachers was the process of annual inspections carried out by provincially appointed superintendents.

The duties of the superintendents in the early Ontario system were to distribute school funds in their area, examine, certify and annul teaching certificates, reporting to the provincial superintendent, visiting the schools to give advice and direction to teachers and trustees as well as preparing and delivering educational lectures to teachers (Houston & Prentice, 1988). Ryerson saw the local superintendents as his "watch dogs" who would control the general principles and character of the school system. These school officials increasingly used measures to ensure that the teachers performed their work as specified in the regulations that outlined their duties (Prentice & Theobald, 1991).
The duties of teachers were increasing dramatically with the addition of a new curriculum that included such diverse topics as morals and patriotism as well as hygiene and callisthenics (Wiebe, 1969). A transformation in the purpose of schooling from learning knowledge to forming morals redefined the relationship between teachers and students (Varga, 1991; Curtis, 1988; Jones & Williamson, 1979). Traditional curriculum and the relative autonomy of the single room schoolhouse teacher were challenged in this age of expansion. The school was now regarded as a means to perpetuate the ethics of early industrial Canada. Teachers were supposed to inculcate their students with the values of obedience, nationalism, punctuality, respect for private property and acceptance of one's place in a competitive economic system (Finkelstein, 1975). This mission resulted in many superintendents trying to model these same values in their dealings with teachers. The focus of educational observers moved from the students to teachers (Varga, 1991).

The *Annual Report* of the Toronto Public School Board of 1886 outlines procedures for ensuring that teachers adhered to new rules and regulations: "Through local institutes, teachers were instructed in matters as personal as their tone of voice and as trivial as how many times to pull the rope when ringing the school bell, as well as in matters more clearly related to academic instruction" (Danylewycz, 1986, p. 70). Qualities such as the teacher's ability to "draw out the forces of his students in an economical and pleasing manner" along with the teacher's "looks, gestures, expression and qualities of voice" were all identified as critical elements of a report (Department of Education, 1855, cited in Curtis, 1988, p. 104). By-law 77(7) of the Toronto Public School Board stated the inspectors were to include in their reports such details as the time they spend in the classroom, the subjects take up and the proficiency of the pupils (Toronto School Board, 1896). Even principals were expected to "examine every class under his charge at least once in each half year" (Toronto School By-Law 111(4), 1896).

Why were teachers subject to such formalized scrutiny? One response is that the increase in enrolment and the lengthening of the period of formal schooling placed teachers in a very important societal role. Teachers could no longer be "without due cultivation or what is worse . . . vulgar low-bred, vicious and intemperate" (Superintendent Charles Duncombe, 1831, cited in Cochrane, p. 130). The amount of autonomy that was to be given over to teachers was a fundamental issue. Teachers were taught their place in this newly emerging system. Educational
experts would recommend the techniques of teaching as well as the appropriate curriculum while superintendents were supposed to monitor schools through observational reports that would lead to the continuation or termination of a teacher's contract. This policy was arguably an attempt to entrench the existing hierarchy of the school system. Teachers were to be reminded at best, or intimidated at worst, to adhere to the rules imposed on them by an administrative structure of trustees, directors and superintendents of education.

School officials at all these levels introduced measures to ensure that teachers would perform their work as specified by the new regulations (Danylewycz & Prentice, 1992). Extracts from reports on the public school inspectors in Ontario in 1872 repeatedly comment on how most of the teachers visited were found "not in conformity" with educational regulations like the limit table and that their teaching methods were lacking:

The question put by the Department, "What method of teaching is employed?" has sorely tried me. In over one-half the Schools there is no method, nor any knowledge of method. . . . The old rote system, if it deserves the name of system, is by far the most prevalent, and so long as this is the case it matters but little what subjects are on the Limit Table. . . . Their pupils prattle learnedly, but know nothing. (Superintendent of the County of Norfold, 1872, cited in Hodgins, p. 157)

The Toronto Normal School was established in 1847, partly in a response by educational authorities to replace questionable teaching methods with standard formal teaching training (Varga, 1991). The focus of teacher training was to include regulating professional behaviour "to implant in them the habits, skills, and the character structure appropriate to the morally forceful teacher" (Curtis, 1988). One of the training texts of the time identified the level of physical control that teachers were expected to exert: "The voice and the eye constitute, unquestionably, fully one-half of the power of a trainer of the young. Thus a trainer's manner may be said to be 'half his fortune'" (Stow, 1854, cited in Curtis, 1988).

J.G. Althouse, a noted Canadian educator and chief director of education for Ontario from 1944 to 1948, claims that the years 1860-1880 saw the greatest advance in the history of school supervision. But these years were also characterized by the voices of frustrated teachers who resented being evaluated by superintendents who had no teaching experience. Superintendents in Ontario were paid 4 dollars per school with the average number of schools usually less than 25
(Althouse, 1929). Only a person with another income, such as a doctor or a member of the clergy, would consider a job with such a negligible salary. Complaints that "crotchety inspectors . . . tyrannize over teachers" (cited in Althouse, 1929) were commonplace. Teachers were expected to "submit silently to the criticism of all" (Stow, 1854). Luckily for many nervous teachers, the reality was that many superintendents did not have the time, energy or compensation required to encourage them to visit many schools, let alone all teachers. Superintendent William Hutton (1888) complained that the twice-a-year visits expected of him were too heavy a demand considering the geographical extent of his jurisdiction (Houston & Prentice, 1988). Urban teachers were inspected more regularly the rural teachers and so were subjected to closer controls. In Toronto, public school teachers found themselves visited by various types of administrators:

... an increasing number of specialists supervised the teaching of subjects like drawing, domestic science and drill. Schoolmistresses and some masters who taught for large urban boards were also visibly compartmentalized in the lower rungs of growing educational bureaucracies which subjected them to several levels of inspection, beginning with the school principal and ending with the district and provincial superintendency. (Prentice & Theobald, p. 145)

Inspectors might have had genuine cause for concern. If the infant Ontario public school system was to survive public scrutiny, teachers had to demonstrate results to their superiors. The inspectors recognized this power. One candidly admitted that, "the teacher makes the school, and the teacher alone" (cited in Stamp, 1982). Robert Stamp's (1982) research revealed that in Ontario in 1876, there were 6,196 teachers (3,405 female, 2,780 male). Less than one-third of them had a first-or second class provincial or county certification. The route to teaching certification was typically to attend the model school where a 14 week course was all that was required before applying for a placement. There was a heavy turnover rate of teachers for a multitude of reasons including marriage or a return to higher paying careers. Only one quarter of Ontario teachers that had graduated from model schools would upgrade their qualifications at a normal school. It was not until 1906 that the normal school in Ontario was abolished and replaced for high school teachers by university affiliated faculties at Queen's and the University of Toronto.

Inspectors that appeared in urban schools were supposed to have a copy of the "limit table" - the regulated time allowance for each subject in every class in the public schools. It was the duty of the inspector to enforce the limit table (Althouse, 1923). Attempts to foil inspectors, who were
sent to find a teacher off topic or out of sync with the regulations, led to the notorious practice of keeping a model lesson handy for the inspector's visit (Fleming, 1971). The possibility that the inspector might show up unannounced at any time kept a majority of teachers in line with curriculum policies. Many inspectors explained that their surprise arrival was due to their desire to see the, "true state of the schools, not what the teachers and trustees wanted to see" (cited in Houston and Prentice, 1988, p. 144). Records show that a few brave male teachers, however, ignored or dismissed the inspector's visit. Robert Galbraith of Orangeville was disturbed that a teacher refused to interrupt a reading class when he was visiting. The teacher in question, when confronted, showed a "violent display of temper" and refused to do what the superintendent asked (cited in Houston & Prentice, 1988, p. 144).

Clara Peterson (1983), an American educational historian, found that teacher evaluation in these years was a very autocraticendeavour: "Superintendents were serving as inspectors rather than helpers: they laid down prescriptions regarding methods and materials, rather than providing teachers with a variety of methods and suggestions and a wealth of materials". Finding fault seemed to be the main purpose of inspection - even the most benevolent of inspectors acknowledged that one of their main functions was to make teachers conform: "His attitude was usually austere and he gave the impression of being there to find something wrong and then to report it" (Robinson, 1972, p. 213). One infamous Toronto high school inspector, Dr. Seath, was noted for "dropping the temperature 10 degrees in any room he entered" (Robinson, 1972, p. 213). A teacher's educational views or divergent classroom experiences did not seem as consequential as his or her superior's need to exercise control.

A typical teacher's contract of the 1870s stipulated the subjects to be taught, procedures to be following and the discipline expected of both the teacher and the pupil. Contracts in Upper Canada commonly outlined duties as the teaching of spelling, reading and writing, arithmetic while maintaining order and suppressing all immoral habits and practices among the pupils (cited in Gidney, 1973). Contracts said nothing about the responsibility of the school inspectors towards the teacher or the school (Prentice & Theobald, 1991). Inspectors at the turn of the century, such as Toronto's Mr. Anglin, would never discuss his report with the teacher he had just inspected. His routine was to arrive five minutes after the lesson began, sit at the back and take notes in his report.
book before leaving five minutes before the end of the class. Teachers were left completely unaware of whether the inspector felt their work was excellent, competent or otherwise (Robinson, 1974).

Naturally there were some compassionate and supportive inspectors. One teacher in northern Ontario recalls how the inspector came in the second week of school and told the teacher to help the new immigrant students live in Canada and not to worry about covering all the curriculum expectations in the limit table (Cochrane, 1981). In their report to the Chief Superintendent in 1873, four inspectors of high schools complained about the effects of centralized control:

> The individuality of some of our best Teachers is repressed, and their energies cramped, or frozen, in the attempt, conscientiously made, to stretch, or contract, their methods to the prescribed form and dimensions. A thorough enthusiast has a more healthy and powerful influence over the youthful mind than the most symmetrical paper Programme ever elaborated; and if great School Masters . . . are ever to be developed among us . . . some play must be allowed to varieties of method, of taste, or intellectual idiosyncrasy. (Hodgins, 1908, p. 239)

There is no doubt when reading the official reports that many of the local public school inspectors, like the principals who would follow them in later years, simply reported what they thought the Chief Superintendent wanted to know, often giving sparse responses to specific questions that they had been asked to answer regarding school deficiencies (Houston & Prentice, 1988). Often the vague comment, "present teacher appears to be doing satisfactory work" (cited in Cochrane, 1981), was the only written record made of the visitation.

It is perhaps not surprising that a founding public school system that was trying to justify its existence by convincing the general population of its need would advocate methods of instruction and administration in both common and grammar schools that stressed order, uniformity and regimentation (Appleton, 1897; Salisbury, 1912). The teacher's delivery of a set curriculum was expected to be a model of systematic precision and efficiency: "The industrial concept of efficiency seems to have been adopted almost intact for use in the field of education" (Peterson, 1982). One teacher, who delivered lessons that included content that was a year ahead of the mandated curriculum, was thwarted by an inspector. When questioned as to why he was not adhering to the regulations, the teacher explained that he felt it was a good idea to give the students
more challenging work. The response from the inspector to this innovative yet nonconforming teacher was very clear: "Don't you do any more thinking; the Education Department is supposed to do the thinking" (cited in Stamp, 1982).

As public education approached the turn of the century, ratings of teachers were commonly used which were based on highly personalized anecdotal commentary made by the inspector. Rating methods, such as checklists were sometimes used to evaluate components of the teaching process (organization of day book, tidiness of classroom and methods of presentation) while the majority of reports consisted of subjective observations noting the discipline, interest and respect for the teacher that students displayed (Peterson, 1982; Varga, 1991). Often the most important item judged in an inspection was how effectively the teacher maintained order in the classroom (Baldwin, 1897; Peterson, 1982). Sometimes, an extremely thorough report was written after the observation:

No defect in the manner, tone of voice, or grammar, is overlooked. Every mispronunciation, error, or defect in stating the successive points of the subject of the lesson, want of picturing out, or failure in securing the attention of the children during these exercises, is plainly expressed. (Stow, p. 350-51)

Comments found in the training register reveal Mr Baker (1887) who made the, "sound of 'n' incorrect", John Forbes (1871-74) who, "speaks low and gruffly", Emma Scott (1883-85) who, "talks too fast . . . voice monotonous" and Miss Weller (1871-74) who "allowed poetry to be read in a sing song manner" (cited in Varga, 1991, p. 22). These comments illustrate that the type of teaching behaviours valued by some rigid administrators. Many of these comments were so trivial, personal and of questionable worthiness that they reveal more about authoritarian excesses that they do about teaching performances. These reports demonstrate how the supervisor was perhaps modeling the same type of power over the teacher as the teacher was supposed to exert on the students.

In the United States, common assumptions about authoritarian supervision were articulated in 1875 by William H. Payne, a superintendent of Michigan schools. He viewed teachers not as professionals, but as workers who followed the economic laws espoused by the rapidly growing industrial society:
It is thus seen that the work of instruction follows the law which prevails in all other industries—differentiation, classification, system; and as in a complicated system of manufacture, while each workman is held responsible for the general results, so in an extended system of instruction there should be a responsible head. (cited in Peterson, 1982, p. 30)

The responsible head was the inspector of schools maintained the authority to tell the teacher how to do his or her job effectively and efficiently as modeled by military and industrial standards. Many Canadian industrialists shared the view that the school must train a child to understand "his economic and social place and responsibilities in our present, highly organized, industrial society" (cited in Stamp, 1982).

By 1869, the Ontario Department of Education formalized observation procedures by publishing guidelines for using Training Registers to observe and evaluate normal school trainees. The guidelines included a lengthy list of characteristics to be evaluated: preparation, fluency, manner, energy, accuracy, watchfulness, mode of posing questions, mode of receiving answers, correction of errors, power of giving explanation, thoroughness, effectiveness and the general value of a lesson (Varga, 1991). An overall rating, from one for "great excellence" to six for "complete failure", was to be entered under each category for each lesson taught. The movement from anecdotal comments to numerical ratings had begun.

Attitudes concerning teacher effectiveness and the role of the supervisor in assuring teacher effectiveness in Ontario were often first articulated and popularized south of the border. Josiah Pickard (1890) wrote a popular American educational book, School Supervision, in which he describes the most important job of the superintendent as monitoring, directing and controlling a teaching force that could not yet be considered as true professionals:

The importance of supervision has grown out of the specialization of labour. Each labourer, becomes through his labour, perfect in the work to which he is assigned, and is thus confined in a narrow field. The labourer knows that his employer expects of him the best results possible for him to attain. He has accepted his place in the plan and he knows it is a small part of a great plan. (cited in Paterson, 1982, p. 32)

This factory model was adapted to the American school system as recognized in Beach and Reinhartz' (1989) study of supervision - students and teachers were often socialized into their
subservient roles. The teacher was the expert on student behaviour and learning while the inspector was the expert on pedagogy and instructional methods.

Teachers, like machinists, were felt to need regular inspection to reinforce the scientific management model. As articulated by the "father of scientific management" Frederick Taylor (1911), teachers, like workers, are "motivated by economics and limited by physiology, needed constant direction" (as quoted in Hoy & Miskel, 1996). In the United States, Boyce (1915) and Bobbitt (1913) claim that they could scientifically appraise teacher effectiveness by observing and measuring teacher behaviour in the class and then applying the time and motion studies of Taylor to them (Adelman, 1989).

The need to provide administrative "experts" with the necessary power and control over workers was analyzed by the American educator and historian, Lawrence Cremin (1961). He characterizes education as a reflection of both the school's presumed expertise and a modern society's rational differentiation of functions. The education "system" should be run as a finely tuned factory whose survival depended on defined power roles. The corporate model of governance, complete with many layers of administrative experts, would keep subordinate teachers "in line" (Tyack, 1987).

The 1870s to the early decades of the twentieth century demonstrate a movement towards a factory model of teacher evaluation throughout North American public school boards (Callahan, 1976; Tyack, 1987). Autocratic inspections were frequently used and perceived as a punitive method to control teachers and to serve as a reminder of their lowly status as a worker. As the Canadian historian W. G. Fleming (1971) observes, although even trained administrators could not reach consistent verdicts about the quality of teacher performance as evidenced in a single lesson, teachers were considered labourers who needed overseeing. Their lowly status was revealed by the belief that an external authority had to maintain centralized educational beliefs through direct inspection and constant monitoring (House, 1989). Some educational historians also view the use of written evaluations as a socialization process that is "deeply penetrating, more total as the surveillance becomes more invisible" (Berstein, 1975, p. 122). Perhaps socialization was the key factor in developing an increasingly standardized form of teacher evaluation. The criteria for what made a good teacher were not yet clearly stated in evaluation policies, but the method of classroom
observation, the existence of regulations like the limit table and the use of written comments were a powerful way to mould teachers into what society demanded of the job.

In these years of early universal education, teaching was not generally viewed as a genuine profession. One Canadian analysis stresses that the influx of young women into the profession was one of the main factors in the resulting devaluing of the work even for experienced males (Prentice & Theobald, 1991). The social views of the day can explain some of the condescension in teaching reports as well as paternal visitations by inspectors to make sure that women were teaching in a competent manner. One teacher lamented in her journal, "I am not sure which inspector thought my long, brown divided riding skirt, made of corduroy, and my spurs, was not a suitable costume for this classroom, but I could not bring a dress, there was no place to change in a one-room school" (cited in Cochrane, p. 129). Ryerson himself at the end of his career had to deal with superintendents who were "troublesome . . . and had little use for what he [the superintendent] considered to be poorly trained 'Female Teachers'" (cited in Houston & Prentice, 1988). In their study of the Ontario school system Susan Houston and Alison Prentice found that local school people complained of superintendents who were prejudiced, who were illiterate, or who based their reports on 'hearsay' rather on 'actual inspection'.

In addition to their classroom duties, teachers were expected to increase their professional knowledge. Inspectors' reports lamented the lack of professional reading teachers did aside from the official Journal of Education (Hodgins, 1908). Danylewycz and Prentice's (1986) study of the emerging school systems of nineteenth century central Canada revealed that during the holidays, after school hours and on weekends, teachers were urged to attend provincially or locally organized classes and institutes, "to learn not only the new subjects but the more modern methods of instruction and classroom management popularized by the 'new education' movement of the period" (p. 287). These extracurricular courses, which were voluntary at first, soon became a standard part of a teacher's workload, further adding to an inspector's potential criteria for what made a competent teacher. This practice of reporting on a teacher's involvement in extracurricular activities, both in student activities as well as those for professional development, has become a problematic part of teacher evaluation which continues in the present. Unpaid labour had now become a routine part of what was considered a valid assessment of teacher effectiveness.
When considering how the public school system fluctuated during its infancy, with its lack of resources, trained teachers, the concern of the public and the radical changes in curriculum and educational law at this time, one can understand why administrators often behaved autocratically. The educational views which called for administrative organization and centralization of curriculum in Ontario schools also demanded uniformity in teacher's duties in the classroom (Manzer, 1994). In an age where the commonly used student texts, the Canadian, Royal and Ontario readers stressed "an omnipotent and omniscient God . . . the virtues of persistence, obedience and truthfulness, (that) idleness and laziness were particularly sinful . . . ." (Stamp, 1982), teachers were expected to model these exemplary traits and maintain the status quo.

Regardless of the push for system-wide control and order, even contemporary journalists claimed that teachers were too rigidly subordinated "to the autocratic fist of an almost irresponsible inspectorate" (cited in Stamp, 1982). Sandiford's (1918) survey of Canadian education at the end of the First World War found that subjects outside the official syllabus could not be taught: "A few enlightened inspectors permit teachers to fit the course of study to local conditions, but this is the exception and against the rule; no latitude in regard to the course of study is the order of the day" (p. 89). A disturbing part of the history of inspection of teachers, is that many of these early attitudes and procedures do not appear to substantially altered in over a century of educational change. Regardless of the occasional inspector who was more humanist in his approach, the main trends of centralized power, externally determined standards and uniformity were frequently found in both administrative attitudes towards teachers and, consequently, in the criteria that were supposed to guide teacher evaluations.

1920s-1950s: The Scientific and Technological Era of Supervision

The next transformation in attitudes and theories of teacher evaluation paralleled the shift in teacher's status from that of labourer to practitioner or professional. This change was due in part to the increase in teachers’ qualifications and training as well as the development of effective teachers’ unions fighting in the political arena for increases in teacher autonomy, rights and pay (Johnson, 1968). After World War I teachers’ anger at their low pay and status motivated
successful union organization. Between 1918-1920, the Federation of Women Teachers, the Teachers’ Association of Ontario, the Ontario Secondary School Teachers’ Federation and the Ontario Public School Men Teachers’ Federation were formed and successfully obtained salary increases, collective bargaining rights and secure tenure for its members (Stamp, 1982). The booming economy of a post-war Canada made it a good time for teachers to flex their collective muscle. Some teachers became more openly critical of the ministry’s focus on patriotic instruction, immigrant assimilation and the checking of subversive political ideas (Stamp, 1982).

Teacher training in all provinces became more demanding. C.E. Phillips' (1957) research into the development of education in Canada found that between 1900 and 1940, the minimum qualifications for admission to training school was raised across Canada to four or five years of secondary education. The Ontario normal school exam now included, along with the history of education, school management and teaching methods, practice teaching reports and the psychology and science of education (Phillips, 1957).

A newly emerging trend in society was a fundamental belief in science, especially in terms of quantitative measurements and categorization as powerful means to “improve” social institutions. Understanding of scientific method included a science-inspired theory of causation that was applied to history, politics, philosophy and economics (Lawr, 1972). It is interesting to note that prior to 1905, no significant research studies had been conducted of teacher evaluation (Peterson, 1982). In sharp contrast, the early twentieth century gave rise to a multitude of studies, especially in the United States, that investigated the application of scientific management and psychological theories on teachers and teaching.

The findings of American educator and psychologist Edward Thorndike (1905-1911) and the French psychologist Alfred Binet (1904) in the early part of the century on assessment of intelligence were increasingly applied to the world of education. North American journals of education at the time dealt almost exclusively with issues of nature-nurture, concepts of intelligence, aptitude and achievement (Church, 1971). The rising prestige of scientific measurement was to have substantial consequences in education. John Dewey, one of the leading contributors to progressivism in education, warned that, "the educational process has two sides—one
psychological and one sociological-and neither can be subordinated to the other, or neglected, without evil results following" (quoted in Fleming, p. 13, 1971). Dewey would have been dismayed at the pervasive focus on the psychological aspects of educational testing at this time.

A campaign began to justify and leadership practices through an application of the new educational science (Brown, 1948; Church, 1971). A burgeoning educational bureaucracy was looking for new theories that would justify the importance of its role. Politicians also turned to educational science for statistical measurements and scientific data to rationalize an ever expanding financial commitment to universal education: "By the 1920s government expenditures on education continued to surpass expenditures on all other social activities and had come to rival government expenditures on economic activities" (Manzer, p. 118). Both business interests and organized labour reorganized costly public education to fit their needs of occupational selection for an industrialized economy. This new focus on class specific and class divided education policies led to more provincial examinations as well as occupational class-structured programs (Manzer, 1994).

In Ontario, the new stress on science and technology led to changes to types of student programs with the introduction and development of separate vocational programs that differentiated between academic and general education. In 1922-23 there were 286 teachers in vocational schools; by 1927-28 there were 660 (Weir & Matthews, 1997, p. 10). The public schools had become more efficient in terms of streaming students that entered the system.

Perhaps, the most questionable reason for the acceptance of "educational science" was to justify the institutionalized placement of children into educational experiences that would sort workers. A competitive system of learning that stressed punctuality, regularity, competition and that rewarded individual merit only helped to legitimizing inequalities among and between students (Finkelstein, 1991). The 1920 School Attendance Act in Ontario generated a system in which all classes of children were expected to be educated. Educational testing, vocational instruction and extracurricular activities were created explicitly to socialize and sort these new types of students, especially immigrant children, who had been brought into the system by egalitarian educational laws (Pennachio, 1986). A teacher was, "expected to supervise a host of new extracurricular activities, offer vocational guidance to his pupils, keep in touch with the burgeoning commercial and business orientation of the period - all in the name of social efficiency and citizen making"
Education in Canada was increasingly class specific and class divided (Manzer, 1994). Did new educational directions effect specific teacher evaluation schemes?

Clara Peterson (1982) found that many American teachers in these decades increasingly resisted evaluations of any type. They saw evaluations as having a three-fold negative effect - as a way to get rid of militant or non-conformist teachers, as an excuse for slashing money from budgets and as a means of enforcing the authoritarian structure of the school. In 1919, Daniel House wrote that no other profession could "parallel the vicious system whereby a self-perpetuating coterie of bureaucrats can play fast and loose with the destinies of thousands of teachers by the simple device of 'ratings'" (quoted in Peterson, 1982, p. 40). As the previous era's methods of industrial management came under attack by teaching federations, a new concept emerged - one that was meant to be based on the pure, objective science of research and experimentation. Teachers found these new concepts much more difficult to collectively debate and resist. What was more, fear of "radical and militant teachers" and their unions may have forced nervous educational managers to embrace scientific principles in an attempt to check growing teacher autonomy (Holly, 1989).

The ratings of teachers based on single classroom inspections were still used but there was a considerable effort in the realm of educational research to find a complex and elusive connection between teacher effectiveness and student progress. The increase in standardized tests of curriculum, aptitude, personality and I.Q. tests for students were to have a corollary effect on the search for new ways to determine teacher effectiveness. No longer were teachers to be inspected. It was felt by many theorists that they too should be evaluated, tested and qualitatively measured (Barr, 1929).

During the late 1920s, intelligence tests became popular in many regions of the United States. Teaching aptitude, subject matter competence and knowledge of testing were also measured (Wilson, 1985). The Cooperative Test Service of the American Council on Education, organized in 1930, provided teacher tests to participating school districts. The tests covered English comprehension and expression, reasoning, general culture, professional information, contemporary affairs and relevant subject areas. Ben Woods, the director of the Council, promoted the use of the test as one criterion for teacher selection, partially due to his belief that a majority of teachers were
mediocre (Wilson, 1985). These areas of testing were the basis of the National Teacher Examinations Program. Such national exams were administered in the United States by Educational Testing Service up to the 1980s.

Some of the more amusing examples of this testing focus in the United States can be found in the newly developed tests of teacher efficiency as advocated by researchers of the 1930s. In 1935, the editor of Kappa Delta Pi research publications, Alfred Hall-Quest, boldly stated that "many important problems in education might be solved if teachers could be classified into qualitative categories on the basis of teaching efficiency" (pp. vii-viii). The editor of the same publication, Helen Walker, advocated both the testing of personality traits and the intelligence of teachers. Looking back on this era, Stephen Gould (1981), a popularizer of biological theories, criticized the 1930s fallacious preoccupation with identifying complicated and multifaceted human traits under the single label of "intelligence". In a similar effort, educational scientists of the Depression era were blinded by a zeal for defining measurable traits that would identify an effective teacher. A barrage of tests was developed for both students and teachers at the beginning and the end of a period of instruction.

In Toronto high schools, psychological and I.Q. testing of select students had been referred to only briefly in Ontario High School's Inspector's Annual Report of 1923, 1927 and 1929 as a way to identify "delinquents" and mentally challenged students but there was no evidence that I could find on the frequency of testing for all students or testing on teachers. However, the American trend towards standardized tests of knowledge and quasi-scientific measured student achievement was evident in Ontario. The development and delivery of the high school curriculum in particular was becoming increasingly dictated by the final examination system (Stamp, 1882). Some specific skills tests were used in an attempt to unlock the mysterious correlation between teacher traits and change in pupil knowledge (Walker, 1935). The use of tests and the scientific focus of education was not just an American phenomenon, as observed by Gordon Sheane: "'Patented techniques' of various kinds were either developed or tested in Canada" (quoted in Tittley, 1990, p. 102).

There are few sources on specific teacher evaluation policies in Canada from the 1930s and 1940s. It does not appear, however, that the American preoccupation with testing teachers was met with
the same enthusiasm in this country. While Canadians divided students into class specific streams of education and used standardized testing and other questionable methods of "counselling" and racial categorization, Canadian teachers may have managed to avoid the prevalent scientific testing that was used on their American counterparts. While American businesses sold their products to many interested administrators, Canadians preferred to only adopt the language and principles of scientific management on their staff (Gillet, 1969).

Education was now often regarded as yet another science, especially south of the border, complete with rules and laws that could be accurately determined and measured. Some of the commonly administered tests given to American teachers both while in training and when tenured were as follows: The Social Intelligence Test (Center for Psychological Service, 1930), The Morris Trait Test (E.H. Morris; Public School Publishing Company, 1928), Aptitude Test for Elementary and High School Teachers (Bathurst, Bureau of Public Personnel Administration, 1927), Vocational Interest Blank (E. K. Strong, Stanford University, 1927), and the Torgerson Diagnostic Teacher Rating Scale of Instructional Activities (Illinois Public School Publishing Company, 1930) (cited in Barr et al., 1935, pp. 82-83). Several attempts were also made to study the influence of intelligence on the relationship between gains in pupil achievement and measures of teacher ability (Walker, 1935, p. 104).

Some of the more questionable measurements of teaching ability were developed by Gilbert Lee Betts, the Director of Educational Research in West Allis, Wisconsin. Tests included the "Test DD", which was developed to measure a teacher's dislike for common activities in teaching such as "meeting socially with members of the community at large" and "enforcing instructions to pupils about safeguarding against contagious diseases." The "Test IA" measured a teacher's acuity in perceiving the relative important of teaching activities. The "Test PP" measured professional pride by comparing a teacher's love of teaching as compared to other possible occupations such as "dairy owner" and "editor of a high class magazine". The "Test RP" was designed to measure a teacher's race prejudice by having him/her rate different races and nationalities by using a one to seven measurement scale--one meaning "might be willing to marry" and seven indicating "would exclude from my country" (cited in Barr et al., 1935).
"Test DS" was thought to accurately measure the extent of a teacher's superstitious feelings. "Test DIS" involved a sight passage in which teachers were to cross out all extraneous and incongruous words as they read. The speed and accuracy in which a teacher plowed through sentences in this particular experiment was meant to gauge teacher distractibility. The educators who developed this test claimed that this activity was "somewhat analogous to the classroom situation in which a teacher is constantly subjected to distracting influences during the process of instruction" (Betts, 1935, p. 169). The following sentence is typical of those found in this dubious "scientific" test: "Even now when there were hot numbers of air children in the same bee class they were vow instructed one hop at a time".

As misguided or overzealous as these American tests may seem, some administrators subjected teachers to them to justify promotion, hiring and dismissal because they believed that these revolutionary tests were objective. Most supervisors between the 1920s and 1950s relied on a multitude of new rating scales to justify actions taken to punish or reward teachers (Peterson, 1982). Who would argue with science and an impressive set of statistics? If teachers wanted to be viewed as professionals, they had to accept more vigorous standards. Many, however, did not.

Strong opposition to the new supervision and evaluation processes combined with other factors such as lower salary scales caused mass protests and resignations in many parts of the United States and Canada (Peterson, 1982). The newly formed provincial federations of teachers in Ontario fought to end what they saw as secretive and punitive evaluation (Robinson, 1972). The inspector still had the power, "at any time to suspend any teacher for neglect of duty, incompetence or misconduct", according to the 1935 Toronto Board by-laws. Yet the unions did manage by 1944 with the passing of the Teaching Profession Act Regulation 18 (1) (b) to demand that any adverse report made against a teacher had to be presented no later that three days after the writing of the report. This was to end the secret and arbitrary nature of teacher dismissals that typically followed a poor performance appraisal by the inspector of schools.

Teaching unions in North America, while gaining many important rights for their members, could not put an end to the profitable business of both student and teacher testing (Gillet, 1969). Due to a combination of public and political pressure that questioned the effectiveness and value of a
growing system, educational theorists and administrators seemed determined to provide many scientific reasons for student placement and teacher effectiveness that would increase measures of accountability. One principal's lament, that was published in the Ontario High School’s Inspector’s Annual Report of 1923, was that there was "an ever increasing tendency to load our public school curricula with purely practical and utilitarian subjects and to make these schools an experimental laboratory for the testing of almost every new fad and frill that science can conceive".

Regardless of the new wave of objective criteria, the autocratic purpose of teacher evaluation can be said to have remained the same. Inspectors in Toronto were to make confidential reports twice a year to the Management Committee to provide the names of those teachers who, "should not be re-engaged by the Board for the following year" (By-Law 97 (7) of the Board of Education for the City of Toronto, 1921). Statements found in the 1920 Report of the Department of Education for Ontario can sum up one prevalent view concerning the relationship between the inspector and the professionalism of teachers:

It is his [Public School Inspector] duty to give such encouragement, direction and advice to the teachers as will enlarge their conception of the work they have to do and to stimulate them with a zeal for doing it . . . . Teachers should be given every encouragement for the display of initiative. It is always to be remembered, however, that unalloyed liberty is, perhaps, the most doubtful of blessings. Teachers are to be found, who, if left to their own devices, would devote the greatest part of their time to some subject in which they are especially interested. To allow such a teacher to follow his inclinations is to place his school under serious handicap. (cited in Gillet, 1969)

Teacher evaluation, based upon scientific principals, would identify many problems with teachers and might even prevent those who might handicap the learning process by "teaching to their strengths".

John Dewey’s progressive ideas about child-centred learning were beginning to take hold in many parts of the United States in the 1930s. He advocated an education to cultivate imagination and to engage children. The classroom was to promote active learning as children constructed, created and inquired. Ontario educators appeared more cautious in their approach. Since the end of World War I, Canadian curriculum and educational practices were generally slow to change. But, by the end of the 1940s traditional practices in Ontario such as the use of school primers, note taking dictation by the teacher, end of term exams, competitive learning environments and enforced
silence were criticized by liberal educational theories (Stamp, 1982). Even so the introduction and gradual increase in psychological tests had also begun to impact on the attitude of the teacher in the Ontario classroom who were observed to have excessive, "zeal for educational statistics, intelligence tests, achievement tests and the like . . . forgetting that [their students] are not pawns . . . but are individuals to be trained and guided" (Weir & Matthews, p. 85). Ironically, some teachers' inability to see their students as individuals was copied in many administrator's attitudes towards those very same teachers. Teachers were considered as another chart to be tabulated, not creative individuals to encourage in terms of their professional growth.

Canadian superintendents continued to visit a minority of classrooms with newly designed rating checklists to guide their written reports (Butler, 1926). Toronto's eight district inspectors' reports show little change in the traditional anecdotal commentary that was used in prior decades to chastise teachers. Most class-room teachers talked too much, "giving the pupils information which they can get for themselves, telling them what they already know, explaining needlessly, and giving unnecessary commands and instructions . . . leaving them helpless when required to think for themselves in new situations" (Stamp, p. 165). There seemed to be little acknowledgement by the inspectors of the teacher's concern with covering the content of the course in preparation for final exams that were not only considered a test of the students but a test of the teacher's effectiveness.

Inspector's reports were considered important documents in a teacher's career, both as a guarantee of continued employment, "my job depended on how well my pupils could recite upon the occasion of the inspector's visit" (cited in Cochrane, p. 72) and as proof of ability to future employers (Cochrane, 1981). In the High School Inspector's report of 1923 on the pupils of Jarvis Collegiate in Toronto, the "character" of each pupil's work was rated through a study of their examination results into charts of statistics with the categories "good, fair, poor, bad". The teachers of these classes were also clearly identified.

Ontario education in the 1920s through to the 1950s has been described as "one of the most dictatorial and thoroughly state-controlled systems of teacher training in the Western world" (Stamp, 1982). Extensive consolidation was occurring from the one room schoolhouse to new
multigrade facilities. Rules and regulations existed for all aspects of educational administration both in and outside of the classroom. Typical of supervisory powers exerted at this time, one of the provincial superintendents, F.S. Rivers, controlled textbook purchases, examinations, budgets and weekly timetables as well as telling teachers which professional conferences they could attend (Stamp, 1982).

In his survey of Canadian education, Sandiford (1918) concludes that the course of study in each province was prescriptive and restrictive and didn't often allow for any modifications from the standardized curriculum: "A few enlightened inspectors permit teachers to fit the course of study to local conditions, but this is the exception against the rule . . ." (cited in Manzer, 1994, p. 89). A few sympathetic remarks can be found concerning the plight of teachers, especially concerning class sizes and the "undermanning of schools". General praise about individual high school staffs were frequently made in annual reports, e.g., "There is an excellent staff, and good work is being done" (1924 High School Inspector's Annual Report, p. 24). One Toronto inspector of the late 1940s and early 1950s, Dr. Hooper, managed to be exceptional in that he won the respect and high regard of the teachers (Robinson, 1972).

The trend of this era was educational researchers seeking to change the criteria in teacher assessment to make it more "scientifically" based. Many administrators were under the impression that checklists and rating cards were a more humane and objective tool to use in teacher assessment (Flower, 1958). In the years after the war, schools became more decentralized with the spiraling school population and building of new schools. It was now principals who had to take on an increasing role in teacher inspections (Candoli, 1995). Simple rating cards or checklists with basic headings for the areas considered important in teacher effectiveness may have been viewed by some overworked principals as a time saving device to use during a brief classroom observation. A sample of five teaching reports studied from one eastern Ontario school board that were dated 1948, 1949, 1955, 1956 and 1957 reveal that these few principals used a more "scientifically" based rating scales with only four criteria (i.e. "preparation of lesson", "teaching techniques", "classroom control" and "knowledge of subject" along with one or two sentences of overall comments.
1960s and 1970s: The Assessment Era of Supervision in Education

It was during the growth of educational testing, standardized curriculum and practical skills-based education that the next tenets of educational reform gained momentum. The human relations theory of management and social systems were accepted by many educators as an alternative approach to the scientific focus of the previous era (Netzer, 1970; Candoli, 1995). Many critics of the educational science movement worked to discredit several tests as being highly unreliable and the results as insignificant (Peterson, 1982). By the late 1960s and early 1970s, the human relations approach was reflected in new discussions about how administrators were to "monitor" and "mentor" teachers (Wilson, 1966; Netzer, 1970). In some cases, teachers found themselves as part of an "assessment process" that considered the complexities of their own personalities as well as the organization for which they worked. Tests and inspections were now frequently viewed as outmoded. Even the title "inspector" was changed in Toronto to "program consultant" (Robinson, 1972).

From the late 1950s to the mid 1970s, as the emphasis in teacher assessment slowly shifted from a top-down autocratic to a more democratic process, the growth and improvement of the teacher was supposed to become the focus of performance assessment. "Evaluation" turned into "assessment" and to improving the multiple factors considered part of pupil learning. Educators felt that valuable change could be achieved through the cooperative involvement of staff members along with administrators who believed in the importance of positive human relationships (Peterson, 1982; Fleming, 1972). The authoritarian models of the past had been partially redefined by democratic concepts.

This human relations movement in education was strengthened by the popular essays of Mary Parker Follett (1941) who asserted that problems in organizations were caused by disharmonious relationships (Hoy, 1996). This human relations method became part of a social science approach that combined theories from psychology, sociology, political science and economics (Hoy, 1996). Mayo and Roethlisberger who conducted the famous Hawthorne studies at the Western Electric...
Company's plant near Chicago (1924-1933) also helped to revolutionize the understanding of the nature of formal supervision. New terms such as "cooperation", "group dynamics" and "decision process" gradually entered and transformed the educational field. Canadian schools remain a hybrid of the traditional and the progressive in regards to teaching techniques as well as attitudes towards supervision. According to Hugh Stevenson's (1970) studies of Canadian public education at the time, progressivism was in the ascendancy.

Thus, previously radical concepts such as open concept classrooms, age appropriate placement and creative student-centered learning became more commonplace. The creation of the Ontario Institute for Studies in Education in 1965 also provided a breakthrough in liberal postgraduate educational studies and research in Canada (Stamp, 1982). The growing complexity of the teaching role, the growth of the media, the increasingly diverse world of science and technology as well as the public's renewed understanding of and desire for a more humanist education for their children are just some of the factors that might have stimulated this new educational paradigm (Wilson, 1966).

Harold G. Spears (1946) had predicted that teacher evaluation would one day concern itself with the improvement of instruction and that attention would be focused on the whole process of teaching and not just the individual teacher, as tradition had dictated (Patterson & Miklos, 1982). The teacher would no longer to be intimidated by inspections or subjected to rating scales. The responsibility for the successes and failures of the education system were now supposed to be shared by an intricate social structure. Superintendents' and principals' leadership skills and style, parenting techniques, biological and social factors in a child's development as well as the teachers' lessons were all equally important in understanding the complexities challenging a child's learning processes. Shared responsibility was the predominant focus in teacher "assessments". Experiments were made in some boards in the realm of assessment to include peers, parents and even students in the process (Netzer, 1970).

Peterson notes that by the 1960s there was more interest from an increasingly educated population for excellence in teaching. Supervisors, superintendents and principals, who had previously been the unquestionable power brokers of the system, were also to be subject to evaluation processes. If
a modern education was to live up to society's new criterion, then the traditional leadership had to be questioned, monitored and improved (Canoli, 1995). Higher standards for certification and continuing professional and staff development became widespread for all teachers (Holly, 1989). By 1968 the largely inadequate practice of summer courses to train university graduates as Ontario teachers was discontinued (Stevenson, 1970). Presumably, teachers were moving away from the role of a learning technician towards the status of a professional capable of criticizing, modifying, researching and developing thoughtful educational theory and practice.

The 1968 Ontario Ministry of Education Hall-Dennis Report, entitled Living and Learning, was widely acclaimed as a revolutionary shift in thinking for teachers and administrators in the system. The progressive authors stressed that both students and teachers should be liberated from the structured, isolating and alienating traditions followed in the education field and both should "learn to emphasize the embracing nature of the learning experience". No more memorizing, mimicking, regurgitating or duplicating in the classroom but instead a new emphasis on the creative nature of the learning process through methods of discovery, exploration and inquiry (Stamp, 1988). In keeping with Canadian's interest in adopting some of the progressive ideas that had been used in many parts of the United States for at least two decades, Living and Learning became a national best seller.

As schools were increasing involved in curriculum development and professional growth, there was a lull in the area of research dealing with teacher evaluation (Wilson, 1966; Peterson, 1983; Beach & Reinhartz, 1989). It is possible that as the status of teaching was slowly elevating from worker to professional, earlier policies of teacher evaluation were viewed as obsolete. It is also likely that research was focused on new curriculum and teaching and learning styles. This wave of new educational ideas resulted in more professional growth. Considerable time and money were invested in opportunities for expanded collaboration. Studies of the 1960s point to a major change towards transferring administrative responsibility away from the central office towards individual schools (Wilson, 1966; Gillet, 1969; Peterson, 1983). The principal became the major evaluator of staff performance and his or her leadership style would largely dictate what type of assessment teachers would face. Decentralized decision making and the focus on individual needs was characteristic of Canadian education at this time (Manzer, 1994).
As superintendents and principals were increasingly held accountable for the public image of their local schools, principals in particular were expected to have expertise in teaching and learning, as well as communication and community relations (Peterson, 1983; Erlandson & Bifano, 1987). Leaders were to establish a climate that would encourage innovation and experimentation. Teachers were given the opportunity to learn and develop new techniques that could not effectively be evaluated by a checklist or rating card. Continuous in-service training became more available - to stress the process of teacher growth rather than a single episode of teacher evaluation.

Increasingly, teachers were expected to participate in their own assessment process. In some jurisdictions they assisted the principal in developing and identifying the criteria of classroom appraisal prior to any visitation (Daresh, 1995). Post conferences became mandated in many policy documents where the principal and teacher, in partnership, outlined successes and suggested areas of growth. An important role of the principal of the 1960s was to keep their teachers up to date with an ever-expanding domain of knowledge through school based professional development. As Robert Goldhammer (1969) suggested:

I have argued that supervision should result in heightened autonomy for teacher, and particularly, in strengthened capacities for independent, objective, self-analysis, and that supervision which increases teacher's dependency upon supervisor to know whether his teaching is good or bad, that is supervision in which supervisor's unexamined value judgments predominate, is bad supervision. (p. 61)

The increasing recognition that the supervisor was a fallible human being instead of a unerring bureaucratic leader denotes a dramatic change in educational attitudes - one that might help teacher assessment become more meaningful, personal and inclusive. Teacher autonomy, not as a threat, but a stimulus to educational improvement, was an essential new belief.

The sweeping social and technological changes of this era were more conducive to progressive attitudes towards teachers as professionals and teacher assessment.

Accountability, evaluation and supervision would acquire new meaning. There would be a shift from the search for centrally administered, objective measures . . . toward independent, qualitative judgments and narrative accounts of experience and performance in learning and teaching. Supervision would concern itself less with monitoring the teacher's coverage of curriculum content than with assessment and
support of the teacher's reflection-in-action. (Schon, 1983)

This new attitude of allowing more collegial and cooperative self and peer assessment was slow to translate into actual policy and supervisory practice. There were critics of this new holistic philosophy to teacher evaluation who predicted the, "protection of incompetence... through pseudo-humanitarianism" (Peterson, 1983). Unfortunately, the critics of a new style of assessment could point to infrequent but damaging cases in which incompetence was protected. This protection probably had little to do with new teacher evaluation techniques - it was due to a complex combination of factors including stronger teachers' unions, the popular trend of personal litigation in any employment dismissal and administrators' unwillingness to be proactive with unfit teachers (Johnson, 1978). Administrators, throughout educational history, have considered critical evaluations as poisoning what could be an otherwise productive working relationship (Coker, 1987). Their decisions on any number of issues, including a teacher evaluation, could now be challenged by the teacher, superintendent, professional federation and legal teams. Administrators may have increasingly strove to avoid conflict and were often reluctant to confront a struggling or incompetent teacher (Bridges, 1986).

Despite these crucial problems, there remained a strong belief that formal evaluation was an essential guarantee that minimal standards would be maintained, even though teachers and principals received little substantial feedback for improvement from these evaluations (Frste, 1994). Once placed in the profession, and past a brief probationary period, only an infinitesimal percentage of teachers were driven out against their will because of poor performance (Fleming, 1971).

It is perhaps somewhat predictable that the participatory concept of teacher assessment would lead to a conservative reaction that became apparent in the late 1970s and continues to gather strength today. A new era of educational accountability witnesses a return to business management techniques of education. Political conservatives and educational traditionalists together question what they saw as the indulgent views of teacher autonomy and professionalism that had been championed in the 1960s.

Some educational theorists link the death of progressive views to an economic downturn (Simons,
Canada found itself in a recession beginning in the mid-1970s, and like other industrialized nations, had a surplus of over-educated workers (Titley, 1990). Governments again increasingly viewed the education system as a potential area for savings and economic regeneration. Politicians "began to favour more targeted and selective forms of intervention and more assertive modes of interaction with teachers to bring the curriculum and its practitioners into line" (Simons, 1989, p. 2). Politicians may have wanted to alter many aspects of progressivism in curriculum; however, boards of education and classroom teachers that had embraced change often withstood this pressure to revert to more traditional methods.

The education system from the late 1970s has suffered from attacks on its effectiveness due partially to economic problems. While in the late 1960s and early 1970s public education reached the peak of its importance in the Canadian public economy, a fundamental change was about to occur: "From the middle 1970s policy priorities turned against education" (Manzer, 1994). The "back to basics" battle cry became one slogan of Canadian politicians and a media responding to the public's concern with fiscal austerity. This concern generated several new policies including more overall teaching time in core subjects especially at the secondary level as well as a new academically "rigorous" core curriculum with less emphasis on options (Titley, 1990). Tom Wells, the Minister of Education in Ontario (1972-1978), introduced a new policy paper that reflected his sentiments that there had been too much autonomy given to schools and teachers in the liberal sixties, and the time was right to return to uniform province-wide evaluations and standard curriculum. Curiously, no substantial evidence of public dissatisfaction with the schools or the teachers was offered to support the Ministry's contention (Titley, 1990).

The optimism of the 1960s in education was replaced by a realistic 1970s: "Little did the liberal educators of the 1960s realize how temporary many of their reforms would be. Even before the end of the decade they were challenged by many fronts" (Stamp, 1988). The rhetoric of business relationships replaced democratic ideals. The 'business administrators' (government) and the 'consumers' (parents) would join forces to hold the 'producers' (teachers) accountable to their 'clients' (students). The education system returned to the idea of payment by results. This was a motive that many educational theorists felt had been abandoned early in the nineteenth century. Public education had become accepted as providing civic education, training for employment, a
means to better one's social class and the route to a cultured mind. The expense of a system that provided so much to its students was accepted by a majority of Canadians, with a few notable critics, by the turn of the century. The "stagflation" experienced in the 1970s probably prompted many people, to question what they saw as a wasteful, mismanaged educational system with questionable standards. The performance of pupils on standardized tests of prescribed curriculum was once again thought to be the only solid evidence of school and teacher effectiveness.

This return to stronger regulated central control was countered by student militancy, parent involvement and teachers' unions. The teachers' federations had grown substantially in membership, monetary strength and political awareness. Teacher salary negotiations resulted in the threat of a province-wide boycott of extracurricular activities. In 1973, a teachers' protest occurred in response to Premier Bill Davies and Education Minister Tom Well's Bill 275 which was a legislated attempt to invalidate the mass resignations of teachers and to enforce compulsory arbitration. Illegal strikes occurred in Windsor and Ottawa. Toronto followed in 1975 with its 8,000 strong membership. This show of force was one factor resulting in salary increases and an increment in professional strength for the teachers of Ontario. The strikes seemed to have no lasting result, however, in halting a return to many conservative educational policies (Stamp, 1982).

There are those who still struggle to maintain the attitudes of the previous liberal era. Many educational theorists assert that teacher development and student development are reciprocally related (Fullan, 1995). Thus, teachers should be nurtured as life long learners who benefit from collaborative professional development, personal and professional reflection, growth plans and honest appraisals of the work they do with students in the classroom. Ernest R. House (1989) aptly summarized progressive attitudes towards teacher evaluation: "Conventional teacher appraisal was oriented towards a labour view of teaching characterized by minimal standards for success . . . evaluation should reflect the more comprehensive view of teaching as a profession and as art" (quoted in Simons, 1989, p. 60). House argued that standards must be set by teaching professionals themselves and that evaluations should be carried out by peers rather than managers.
Contemporary Teacher Performance Assessment

Educational administration theories of the 1990s detail many aspects of supervision that were intended to result in changes which included teacher evaluation policies and procedures. The focus of substantial analysis in the 1990s targets the increase in bureaucratic systems being applied to educational systems. A fundamental cause of conflict for teaching professionals comes from being subjected to the systems of social control used by bureaucracies. Many administrative theories advocate for education to accommodate this conflict by establishing loose structures or developing dual authority systems (Hoy & Miskel, 1996). And yet in the realm of teacher assessment the principal, or more frequently, the vice principal, remains as the sole dispensor of rewards in this system. The classroom observation and the written report are two of the means of social control which is followed by a third - what Donald Haefele (1993) labeled the “tell and sell” of feedback. When teacher and principal meet to discuss the written report, the principal defends the process, decisions and recommendations: “This is a one-way communication process from the principal to the teacher . . . a monological process (Haefele, p. 336). Principals dominate and the teacher remains passive in this bureaucratic process of control: “The teacher wants to look good and conceal, albeit deny, any weaknesses. The principal wants to suppress objections and command respect for his/her authority and judgement” (ibid). Teacher evaluation corrects and controls teacher behaviour in specific ways, namely to achieve organizational goals.

Conflicts have arisen because of the duel expectations on teachers to act in the best interests of their “clients” and the “organization” (Hoy & Miskel, 1996). Teachers, as professionals, act in a manner that benefits their students. The criteria and mechanisms for evaluating teacher performance has become more managerial in that a board defines its mission, goals and objectives that its employees are expected to promote and achieve with some measure of efficiency. One Board’s evaluation policy stated that the procedures of teacher assessment were aimed at developing “a framework that would guarantee system coherence” (Growing Together, 1992). Control in bureaucratic organizations stem from one major line of authority. In contrast, professionals are given autonomy in deciding what is best practice and at times may be censured by
colleagues. Self-imposed standards and peer-group surveillance may be the desire of most teachers. Yet teacher reports have become a bureaucratic formality. Hoy and Miskel (1996) seem correct in their assessment of the increased frequency and extent of the conflict between the bureaucratization of schools and the growing professionalization of teachers.

What I defined earlier as the peculiar combination of top-down bureaucratic inspection masquerading in humanist terms in the 1990s can be attributed to goal-setting theory being widely supported within organizational science (Hoy & Miskel, 1996). Teacher behaviour, like the behaviour of employees in other business organizations, is thought to be regulated and maintained by goals. In the most current practices in teacher evaluation in Canada and the United States, a growing part of the written report is devoted to a teacher’s “growth plan” or goal-setting (Fraser & Streshly, 1994; Candoli, 1995;). Phrases highlighted from one Board’s document on teacher assessment reveal humanist and progressive vocabulary: “. . . unique needs, talent and abilities should be nurtured”; “A teacher’s positive image should be considered. . .”, “Teachers should be encouraged to be life-long learners”; and “Cooperation and dialogue between teachers and administrators enhances the educational process” (Growing Together, 1994). These examples are typical of the language but not necessarily of the intent of teacher evaluation in the 1990s as will be argued below. A “one-size-fits-all” hierarchical model remains (Haefele, 1993). One hundred years later I find that little has changed from the comments made by one Ontario Minister of Education, who admitted in 1892 that teacher inspection was "more for examination than instruction" (cited in Phillips, 1957).

Taking a more focused look at the problems associated with actual writing of a single report was not a topic of any widely published educational research until the 1980s and 1990s. Lawrence Rudner’s (1992) work was important for highlighting the problems and errors in assessment procedures that involved pure observation regardless of pre determined criteria. These errors greatly reduce the validity of scores. He categorizes five main problems. The first was the "halo effect" that occurs when the person being rated is not performing in their usual manner, the situation may not elicit typical behaviour or the evaluators may unintentionally distort the results. Rudner (1992) found that Nisbett and Wilson's 1977 study of student rankings of a professor based on videotaped lectures confirmed the occurrence of the "halo effect". In one tape the professor
acted in a friendly manner, in the other he behaved arrogantly. The students rated the friendly professor higher in traits such as physical appearance, knowledge and mannerisms. A variety of traits such as the physical appearance, age, gender or ethnic group of the teacher would influence the report of the nineteenth century inspector as it does his or her twentieth century counterpart.

The second problem that Rudner identifies is stereotyping. The impressions that an evaluator forms about an entire group can affect how he or she interprets behaviour. For example, an inspector might rate a science teacher as exacting and fastidious because he or she feels that all science teachers are supposed to act in this manner, whereas, a drama teacher should be more creative and spontaneous.

The third problem is perception differences. The viewpoints and past experiences of an evaluator can affect how he or she interprets behaviour. A principal that has never taught a certain subject area has different perceptions from those of a principal whose experience includes 20 years in the same subject area as the lesson observed.

The last two problems, according to Rudner, involve statistical validity. "Leniency and stringency" errors occur when a judge doesn't have enough knowledge to make an objective rating. He or she may compensate by giving scores that are systematically higher or lower. Again, the experience of the person making the evaluation will lead to an increase or decrease in validity. The last problem Rudner identified was "scale shrinking" which occurs when the evaluator will not use the end of any scale. An administrator who takes pride in never awarding a teacher with an excellent rating is just as unreliable as an evaluator who would never want to report a failing performance.

The problems that Rudner outlined in the 1990s can be found in the one eastern Ontario urban school board's teacher assessment policy and procedures that I studied because of the reliance on a single classroom observation as the fundamental means to make evaluative judgments. While an established body of literature (Latham, Wexley & Purcell, 1975; Jaeger & Busch, 1984; Pulakos, 1986) argues that extensive training can minimize these numerous threats to the validity of observational evaluation, this type of training was non existent until the 1960s, and in the majority of cases, as will be discussed later in this paper, is still not routinely offered.
Alternative methods of teacher assessment are plentiful in the literature of educational theory and research and can be found in pockets of progressive schools in the U.S. and Canada, but they have yet to be widely adopted (Haefele, 1993; Skinkfield, 1994). While educational theorists have urged administrators to find more opportunities for evaluation procedures in which the evaluator becomes a monitor and not just a judge of a teacher’s work, Daresh and Playko’s (1995) educational studies have shown that this practice still does not exist in many North American schools.

Teachers in the late 1980s and early 1990s have expressed a preference for having different sources of information in their assessment in addition to classroom observation such as self-evaluation questionnaires, portfolios, peer coaching and reviews - but these, also, are not frequently implemented (Lawton, 1988; Airasian & Guillickson, 1993). The use of standardised (anonymous) questionnaires to get student feedback especially at the secondary level could also be another method of promoting positive change (Scriven, 1994). The use of multiple sources of information regarding teacher performance have been promoted in educational literature of the 1990s as one crucial way of improving the assessment process and making it more than just another administrative necessity.

The newest terminology to emerge in the theoretical considerations of teacher evaluation policies is the "review" or "performance assessment". The review is a model found frequently in businesses and corporations in the 1990s. There are two basic types of review. One is the summative review, which is purely coercive, that is used for purposes of hiring and firing. The second is the formative review that is used solely to identify strengths and weaknesses of the employee and to assist him or her in making improvements in preparation for the summative review. "Management by objectives", "management by results" and "clinical supervision" are included in the new lexicon of performance appraisal in education which reflects trends started in the 1980s in business and government (Musella, 1988). These models include a basic management principle which advocates a three stage process to teacher evaluation: appraisal, support and continued professional growth (Iwanicki, 1990).
Distinguishing between these two very different purposes has yet to happen in most school board policies: "Supervisory and administrative personnel have not traditionally worked with teachers in formative ways" (Daresh & Playko, 1995). The majority of board contracts in Ontario demand that all teachers have a written evaluation based on a classroom observation, every two to three years (Lawton, 1988). Classroom observations are the most common data sources influencing evaluations of instruction in today's schools (Cangelosi, 1991). Probationary teachers are evaluated every year and the results determine if a permanent contract will be granted. The typical board policy of a 40 minute classroom observation for elementary teachers and a 70 minute observation for secondary teachers, once every two to three years, that is based on nebulous criteria, cannot yield reliable data to determine effective teaching (Musella, 1988).

Literature on Assessment Policies and Procedures in Ontario

One study into teacher assessment policies and procedures was based in Ontario in 1984 (Lawton et al.). Funded by the Ministry of Education, the study dealt with four main areas of concern - types of policies, the extent of implementation of policies, effectiveness of appraisal practices and the processes used to develop and implement policies. The overall findings of the study included the point that more effective appraisal systems held pre-conferences, had clear communication of purposes, used general and specific notes, reported under several headings and involved post-conferences as well as the making of plans. Respondents to their survey reported that they preferred that more sources of data be used in their appraisals and that while groups approved self-evaluation questionnaires, they were rarely used in practice. Teachers in the public system did not support criteria that were related to out-of-classroom activities while teachers in the separate system were more supportive. They also found that appraisal systems that were reported to be effective have the least incidence of placing staff under review of dismissing staff or filing of appeals, grievances and lawsuits. Whether this was due to the supportive or lenient nature of the appraisal system cannot be determined from the data collected.

Another conclusion from this report noted a disposition by teachers towards the evaluation process - it had no demonstrable effect on them: "Only a very small percentage reported any positive rewards from a positive evaluation, and an even smaller, almost negligible, number reported
negative effects" (Lawton et al., 1988, p. 29). The authors of the study found that, "improvement of performance was not commensurate with the amount of effort put into evaluation. At best, we noted that only rarely did respondents feel that appraisal had any negative effect on their performance" (p. 32). The teachers and administrators I have worked with frequently comment on the investment of time into the procedure. Many have stated that they wish that the evaluator would just "show up" and write a report. These few people cannot see any value in conferences, discussions and debriefings with their administrator, perhaps because of the lack of obvious repercussions - positive or negative in nature.

This extensive study of 187 Ontario school boards found that while policies reported that the primary purpose of appraisal was developmental, only 14 percent of the teachers in the study reported any kind of plan resulting from evaluation. Only half of those teachers that reported a plan indicated that it was monitored in any way.

Richard Stiggins' study of teacher evaluation in Portland Oregon in 1986 was also critical of the process and led to questions regarding its usefulness:

For 99.9 percent of teachers, all of whom are at least minimally competent, evaluation results (if gathered at all) are compiled, entered on the standard form, signed and filed away, never to be seen again. Requirements of the law and the contract have been used up. Both teacher and principal have done their job. The myth of rigorous personal management remains intact. And the evaluation process has helped almost no one to improve. (Stiggins, 1988, p. 141)

Only extensive current research on Ontario teachers and administrators' attitudes will reveal if there has been any fundamental change in these problematic incriminations.

Teacher Performance Assessment in One Ontario Public School Board

The rating scales and other types of "pseudo-meaningful instruments" (Cangelosi, 1991) that are used in some American schools were not used in the particular Board that I examined. In order to analyze the connection between explicitly stated Board policy on teacher evaluation and the actual written documents, I was granted permission to examine 20 teacher reports that are kept in teachers' personnel files. I wanted to analyze the reported purposes of teacher evaluation to see if
they were actually being achieved. I also wanted to compare the contemporary purposes and techniques to the four historical trends to assess what, if any, fundamental changes had occurred. Through opportunistic interviews with several colleagues in the past ten years of my employment with this Board, I have also collected interesting personal insights into the policies, procedures and effects that teacher evaluation has had in five specific secondary schools through the leadership of eight principals.

The reports analyzed are, in theory, important documents to all teachers. They have to be signed and dated by the principal and teacher; the document is then forwarded to the superintendent's office for final approval. It proceeds on to the Human Resources department to be filed in the teacher's personnel records. These reports may be especially important for teachers who are seeking promotion. Postings for positions of responsibility will explicitly request the most current teacher report to accompany a resume. From my limited experience on seven interviewing panels at the high school level, I have witnessed that a less than glowing report may stall a teacher's career aspirations. At the very least, the reports can also be crucial for those teachers on probationary contracts. Probationary teachers in this one urban school board in eastern Ontario must achieve a satisfactory rating in order to be granted a contract.

Aside from these two groups of teachers who are most directly affected by the results of their teacher report, all teachers are observed and evaluated every two to three years. The Board's policy states that three specific purposes of evaluation exist. These purposes are to help teachers improve, to identify their strengths and weaknesses and to propose growth plans for implementation. These are pedagogically explicit reasons that reflect the historically pervasive belief in teacher evaluation as one way to better education. The reports that were written would presumably reflect some of these noble aspirations. Unfortunately, the reports I examined revealed fundamental contradictions between the purposes of teacher evaluation and the actual statements made in the written reports.

The 20 teacher reports were pulled from randomly selected files held at the Board office. Access to these files is only permitted to a teacher who requests to view their own file and to supervisors who wish to read them. Two supervisors in this Board said that files are read only in cases of misconduct and when teachers are placed “under review” to see if a pattern of documentation has
been established by administrators regarding unprofessional conduct and or neglect of duty. The most frequent request of teachers for access to the files is to photocopy teaching reports to submit with a resume for an interview or other job opportunity. Since these files are confidential, I completed an ethical review as well as discussing the purposes of my research with the supervisory officer in charge of the department of human resources. He admitted that such a request had never been made in his recollection, and added that “not much has changed in teacher evaluation since I’ve been here”. My request for access was granted and a letter was signed by the superintendent stating the purpose of the research and the methods I would use to guarantee the anonymity of the teachers, principals and superintendents. I requested that the files be selected at random by the administrative assistant who monitors the room that stores this information. Half of the files belonged to teachers in the elementary panel; the other half belonged to secondary school teachers. The teacher report examined was the most recent in the teacher's file. The oldest report was from 1988, with the majority of reports dating from 1996 to 1998. To secure confidentiality, all names and identifying information were removed from the documents. I returned one file in exchange for another when I recognized a name of a colleague who had written the report. The reports were photocopied in the file room and then returned to their rightful place. After completing an analysis of these recent reports, I did return to examine five archival files of retired teachers who were assessed in the 1940s and 1950s to reaffirm trends that I identified in that particular time in education.

With the reports and the policy documents as the main sources, I will offer a summary of the Board policy in several different areas such as purpose, criteria and guidelines to classroom assessment. I will critique the Board’s stated policies through additional reviews of current research on teacher assessment as well as synthesizing the written reports in terms of the degree of success that Board policies, however flawed, were followed. The reports revealed other interesting trends in a study of the vocabulary used in the descriptions, the differences in the reports of male and female teachers, the person who was responsible for the evaluation of the teacher (i.e., principal or department head), the written accounts of pre and post conferences as well as the actual length of the written reports. I will comment on what these samples may reveal about the complex process of teacher evaluation based on a written summary of one classroom observation. I will suggest alternative methods of teacher assessment or modifications to the practice followed in this Board
that are found in current educational administration theories.


Since the reports spanned ten years of evaluation policy in this Board, they reveal important changes in one decade in reporting techniques and formats. In 1982, the purpose of teacher evaluation was stated on a one page guideline sheet: "The purpose of teacher evaluation is to improve instruction and to encourage the professional growth of all staff". This guiding purpose is problematic. It combines both formative and summative goals in assessment. Formative assessment is to increase teacher effectiveness in the classroom whereas summative assessment is used to promote, hire and fire teachers (Caugeloni, Stiggins & Bridgeford, 1984).

Once a reported purpose is articulated concerning the reasons for teacher evaluation, educational administrators must develop what they believe to be necessary criteria for the evaluations. Back in 1896, John Millar, then the Deputy Minister of Education in Ontario, stated his criteria for an effective teacher: "A good teacher should have physical health and vigour, scholarship, professional attainments, personal magnetism, executive ability, tact, common sense, vigilance, heart power, will power and moral character" (cited in Phillips, 1957). The 1950 Royal Commission on Education in Ontario listed the qualities of a teacher that were most valued at that time as scholarship, professional skill, "a high standard of physical and mental health, superior intelligence, a deep and abiding religious faith, a mature and stable personality, and a willingness and ability to mingle and co-operate with his fellows . . . a general culture . . . [and] aesthetics to appreciate beauty". The criteria in this Board from 1982 to 1990 are quite similar in intent to those used a century or even 50 years ago.

This is the itemized list of criteria that states possible indicators for the assessment:

The following are some observable indicators of effective teaching: warmth, a caring attitude towards students, orderliness and organization, flexibility, responsiveness to individuals, knowledge of subject matter, clarity of presentation, encouragement of student initiative and independent thinking, classroom control, use of student answers in the development of the lesson, enthusiasm, absence of negative criticism, and variety in the use of instructional materials and techniques.
As I will argue below, this list is fundamentally flawed with criteria that the Board has falsely deemed "observable indicators of effective teaching".

My stating point for this critique is that the criteria for evaluating instruction should focus on the quality, value, effectiveness and impact of the instructional activities undertaken in the classroom, not on the perceived value of the teacher themselves (Cangelosi, 1991). No matter how it hurts our educational sensibilities, a teacher does not have to outwardly demonstrate a caring attitude towards students or be enthusiastic to have their students learn. While these personality traits might make the classroom a more enjoyable place to be, they are not reliable, valid or useful indicators of increased student learning. Research findings do not support the entrenched belief that there is a relationship between student learning and a teacher's sense of humour or rigid classroom management style or any other similar teacher attributes (Cangelosi, 1991). In the end, the evaluator should be focused not on how to change the teacher's personality but on what and how the teacher teaches in order to improve instruction for the students (Daresh & Playko, 1995).

The Board's "Teacher Evaluation Guideline" continues by listing "additional criteria which may be used as discussion guidelines: professionalism, contribution to the total school program, interpersonal relationships, learning environment, quality of program, student assignments, evaluation of student progress, realization of objectives". Again these additional criteria range from the specific (e.g., student assignments) to the indefinable (e.g., interpersonal relationships).

To start to take apart those points, there is no description in the evaluation guideline to define a good learning environment. The Teacher Evaluation Guideline itself is a mere one page in length. The guideline states that the additional criteria are for "discussion", meaning that none of these criteria should be found in the written report and that only the "observable" indicators should be included which is a positive point considering interpersonal relationships and evaluation of student progress could only be assessed in ongoing monitoring. In effect then, the teacher has one class period to demonstrate his or her knowledge of subject matter, classroom control, enthusiasm, warmth, orderliness and intellectual grasp of the subject area. Fundamental change in policy or practice since the inspector's visits in the early part of this century has not occurred in this Board.
This traditional reliance on subjective observation is somewhat tempered by a list of indicators to be used as a guide by the supervisor. From the Board's list of observable indicators that would be used in the written report, 6 of the 12 items are arguably personal behaviours (warmth, flexibility, enthusiasm, caring, responsiveness and absence of negative criticism). These are hardly completely reliable measures of good instruction that are intended to measure what and how students learn as well as what and how a teacher learns professionally. Persuasive data should contain no value judgments, inferences or generalizations that an evaluator might have formed during a classroom visitation (Stodolsky, 1984). While this may be difficult to achieve in practice, the additional criteria found in this Board document are not only focused on the teacher's interpersonal behaviours but they are so broad that they require explanations, definitions and parameters. Discussing a teacher's "professionalism, interpersonal relationships, and realization of objectives" would be an insurmountable task without valid and measurable indicators of these traits. To achieve the reported formative and summative purposes of the written report, the criteria used for judging teacher effectiveness must refer to specific job responsibilities, not just vague subjective descriptors of the teacher's personality, attitudes and values.

The Recommended Process of Evaluation

The teacher evaluation guideline includes a summary of the process of the teacher evaluation.

The Principal should meet with the staff to discuss the evaluation program and the criteria to be used in the evaluation process. Each school, because of its particular nature and mutually developed goals, may add to the list of criteria. The evaluation will be done by one or more of the following: Principal, Vice-Principal, Department Head or Area Superintendent. A discussion with the teacher should be part of each evaluation.

The use of the word "should" is disconcerting. This document states that while principals are not required to discuss evaluation criteria, they could add to the criteria as they deemed appropriate, and that they did not even have to discuss the evaluation with the teacher. This is a surprisingly explicit statement to make. Common sense and sound educational policies would dictate that a teacher being evaluated must know the criteria by which they are being measured. There were several documented cases in Ontario in the late 1970s and 1980s of unpleasant legal repercussions.
following the problem of teachers, who had received a negative reports, claiming to not have known the criteria by which they were judged (Hickcox, 1988). Yet principals in the late 1980s, according to this written Board policy, still did not have to discuss the criteria of evaluation with their teachers. Here lies the potential for a completely summative, sometimes punitive "top-down" exercise. It can be argued that teachers as professionals should be responsible enough to know any possible criteria. Various sources such as educational journals, board policy documents as well as College of Teachers guidelines provide lists and descriptions of currently acceptable criteria.

However, the principal could even "add to the list of criteria" developed by this Board policy. This document from 1988 hearkens back to the evaluation techniques of the early part of our educational system, notably the inspection of teachers in an arbitrary manner by administrators armed with a copy of the Limit Table and regulations who would use questionable criteria chosen completely by personal preferences. In a response to the problem of teacher awareness of how they were being assessed, the Ontario Secondary Schools Teachers' Federation has repeatedly issued a "performance appraisal system evaluation chart" to all its members which states that the purposes, objectives and criteria chosen must be clearly defined (OSSTF, 1980, 1985, 1990, 1993).

A department head in this Board may also be called on to conduct the evaluation of a teacher. Department heads in Ontario secondary schools gain their positions by having a specialist degree in the subject area in which they wish to become a leader and by being successful in an interview process. Department heads have not had any specific training in evaluation techniques that should be part of the training to attain principal's qualifications. The role of the department head has traditionally been to timetable subjects, maintain a budget, lead in the development of subject specific curriculum and model good teaching strategies and evaluation techniques. They had never, until the late 1980s, been asked to evaluate members of their department. Heads in the public school system were seen as mentors, not reviewers of the teachers in their department.

Under the Education Act (1980), evaluation of teachers is not defined as a task for department heads. Principals, according to the same Act, can assign specific duties to any teacher they wish. Rumulo Magsino (1987), in his examination of teachers' legal rights, states Canadian education acts serve to support the common-law tradition which defines the school board-teacher relationship
as a master-servant or employer-employee relationship, regardless of teachers' perceptions of their academic and civil liberties. Whether it is wise or prudent to have department heads evaluate other teachers has not been thoughtfully considered by Board administration. The reality of the massive cost in time and energy for principals to regularly evaluate their staff has resulted in a new trend. Department heads in this Board and throughout Ontario can now be asked to evaluate department members, regardless of a multitude of problems: they have had no specific training; their perceived role is as a supportive coach and ally, not an evaluator; and the Ontario Secondary School Teachers' Federation specifies that department heads should refuse to take on a role as official evaluator (OSSTF, 1980, 1993). I have witnessed many heads' personal and professional discomfort in participating in this summative type of evaluation. In this connection, I know of colleagues whose principal stated that department heads would observe and write a report on all members of their department that were due for the official evaluation, unless, the teacher was "at risk" or had received numerous student or parent complaints. In these "tough" cases, the principal would relieve the heads of this difficult duty. The reality of teacher evaluations is that, for many administrators, at least in the five high schools that I have been employed, it has become yet another tiresome duty that does not add to organization effectiveness or to the school's educational goals.

Of the 20 teacher evaluation reports studied, the difference in who is responsible for making the report seemed to depend largely on whether the teacher taught in an elementary or a secondary school. Of the 10 elementary reports, principals wrote 8 and vice principals wrote 2. In the secondary panel vice principals wrote 6 of 10, department heads wrote 3 and a superintendent wrote 1. The report written by the superintendent was unique in that the person being evaluated was looking for a promotion to vice principal. When it comes to the critical task of choosing potential candidates for promotion, the classroom observation becomes more of a concern for an administrator. None of the secondary evaluations, it turns out, were written by principals. The reports may confirm the tendency in the past two decades in Ontario of principals assigning some of their evaluation duties, in secondary schools, to department heads (Hickcox, 1988). One could argue that principals should take very seriously their role in improving instruction and encouraging professional growth but it would seem in the sample studied that at the present time, observing teachers' classes and writing formal reports is not considered an important or practical way to
achieve these educational goals.

An interesting comparison can be made when reading the secondary evaluations that were completed by vice principals versus those written by department heads. As might be expected, the department head commented much more frequently on the teacher's command of the subject area and his or her ability to work with other members of the department. The vice principals, however, presumably might not have the subject expertise; therefore, they did not comment on subject specific content, but focused instead on generic teaching skills (e.g., evaluation techniques, classroom management) in addition to the teacher's commitment to extracurricular activities in the school as a whole. Comments such as "Ms D. prepares clear, detailed, attractive handouts . . . ." and "The lesson involved the whole class" were typical of the observations made by vice-principals. One report written by a vice principal demonstrates the tendency to focus on performance outside of the classroom:

Mrs. R. is a team builder. She constantly encourages staff to get involved and assists them with programming. She acts as a liason between staff and administration bringing their concerns forward as well as offering possible solutions. She also deals with parents most effectively, pointing out to them when the need arises, the positive aspects of their child's educational experiences . . . . Mrs. R. is very adept at dealing with concerns of students, staff and parents. She deals fairly and consistently with such matters in a way that is understood and accepted by those involved. Mrs. R. has been an active member of our School Management Team . . . . She has coordinated a number of school initiatives on special occasions . . . She has coached . . . She assists with out weekly Administration Meetings . . . . She was involved as a member of our school team . . . . She is a valued member of School C's team.

Only one sentence of the entire three page report on Mrs. R mentioned her teaching skills: "Mrs R. is an excellent classroom teacher." It seems questionable to me whether or not classroom performance was actually observed at all. This focus of this evaluation would seem more appropriate for a chairperson, a department head or even an administrator, not for a Grade 4 teacher.

Department heads, on the other hand, noted subject based skills such as French and math with great detail:
As the students walked into the room, they were greeted with music and news reports from the French radio station. When the class was quickly settled and ready to start, volunteers put the answers to homework previously assigned on the board. Homework questions reflected the theme of the present unit "Bien dans sa peau" (feeling good about yourself) and Mrs. M.'s questions generated excellent answers which showed not only a mature thought process but also a very high degree of proficiency in the French language. Mrs. M. took the opportunity to review language structures already learned and concentrated on grammatical points . . . Mrs. M's excellent command of the French language, her clear and precise wording, her encouragement and praise for their efforts . . . generated maximum participation from the entire class . . .

In the class visited, the objective was to consolidate graphing skills, and learn to use the TI-81 Graphics Calculator for graphing functions. Mr. S. emphasized and reviewed various skills and properties that the students should understand: the equation has the same graph no matter how it is rearranged; how to rearrange equations to solve for x or y . . . The students enjoyed trying to create interesting graphs and figuring out how to make them.

In sum, the evaluator significantly constructed what was being observed.

Problems arise when administrators who have no subject expertise are expected to evaluate teachers by using several criteria under the heading "subject/program expertise". Being evaluated by a department head might ensure more reliable data regarding subject expertise, but the department head, and the principal for that matter too, may lack the training and knowledge of current educational theories. Peer review has been adopted in many professions outside teaching with recognized results (Wise & Darlene-Hammond, 1988) but remains problematic to implement in schools due to the entrenched formal, structured and standardized function of current teacher evaluation policy. Both the subject expert and the educational leadership expert should be called on to develop criteria specific to each teacher and to perhaps cooperatively write the report and mentor the teacher on ways to improve.

The final section of the "Teacher Evaluation Guideline" states that general assessment ratings were to be "used for all probationary teachers and for permanent teachers whose work is unsatisfactory or poor". The general assessment ratings are as follows:

Excellent: An outstanding teacher whose performance ranks high in all respects.  
Very Good: A teacher whose performance is highly commendable.  
Good: A teacher whose performance is without any major weakness.
**Unsatisfactory:** A teacher whose performance is below an acceptable standard. An unsatisfactory teacher must show improvement in those areas outlined in the evaluation report.

**Poor:** A teacher whose performance has major defects. A poor teacher must take immediate steps to correct these major defects, failing which a recommendation will be made to the Board for termination of the teacher's contract.

This rating scale is illuminating. After the classroom observation and the writing of anecdotal comments in a report, the principal has to choose one of these overall ratings for every teacher. The entire evaluation process is tainted in its purpose since only probationary teachers and permanent teachers whose work is unsatisfactory or poor must receive a rating. Deeming a permanent teacher's work as excellent, very good, or good was not viewed as necessary. It would appear that the overall rating serves as a punitive function to "flag" incompetent teachers and to indicate whether a probationary teacher should be granted a permanent contract. This is purely summative evaluation. There is no desire to improve a teacher's performance - only the threat that if they don't improve by the time of the next evaluation, that his or her employment is at risk. In my informal discussions with colleagues in the secondary system over the past ten years I've noted teacher satisfaction or dissatisfaction with an evaluation report having less to do with the written comments and more to do with the one or two word rating. Why a teacher was "good" and not "very good" or "very good" and not "excellent" could not be easily justified by the administrator and in some unfortunate cases may have caused strained relations between the teacher and the evaluator.

The descriptors used in the general assessments are, in many ways, lacking in clarity, reliability and validity. The key word used in the assessment is "performance". If the "performance" in one particular lesson is "highly commendable", than the teacher is excellent. There is no attempt to define an acceptable, let alone exemplary, standard by which the teachers are being measured.

1990 Evaluation Policies and Purposes

In the Board where the reports were reviewed, the "Teacher Evaluation Guidelines" were updated in 1990. What major changes were made to the written policy to assist the evaluator and the
teacher? The first section of the guideline outlines the format of the report:

Components of the Teacher Report-

a) information and signatures specified on the report,
b) observation,
c) areas for growth, and
d) general assessment

When reading the 20 written reports, the majority of the content was devoted to a description of the teacher's lesson - on average between 80-95 percent of the sentences. Often several paragraphs, if not pages of writing, were observation notes while the "areas for growth" section received two or three sentences:

Mr. Q. is currently teacher Math and Physical Education at . . . . This lesson observed was a Grade 9 Physical Education class in soccer. The students arrived promptly in full equipment ready to get going. They obviously new what was expected by Mr. Q. and were looking forward to the class. He spent a few minutes at the start of class to return written assignments and gave them an opportunity to return them to the change room. This was followed by a full body warmup activity in the gym before going outside and running the mandatory 800 metres. He then moved into a series of soccer drills . . . . He discussed some soccer strategies . . . . Overall the lesson was well prepared, and offered much in the areas of soccer skill practice and fitness. . . . Mr. Q. should continue to seek opportunities to visit other classroom teachers when possible, to compare what he is doing in the classroom, to find new ideas and to share his experience with other teachers for the benefit of everyone.

This was the norm regardless of the general assessment that the teacher "earned". If one of the fundamental purposes of classroom observation is to improve teacher performance, than a teacher who was deemed merely "satisfactory" should be a prompt to the evaluator to take considerable time in outlining a growth plan for this teacher. Yet in my experiences of evaluation, which are shared with my colleagues at the same schools, most principals request that a teacher submit a written growth plan before the classroom observation has even occurred. This growth plan, then, is often copied word for word into the written report.

While it certainly may be beneficial for teachers to develop their own professional growth plan, the
argument can be made that this process should be balanced with the Board's other stated purpose of evaluation which is to improve instruction. This process would be facilitated by the principal, or "master teacher" discussing, defining and advising collaboratively with the teacher concerning suitable areas for improvement. A teacher who identifies computer literacy as one of their main areas for professional growth, when it is clear that classroom management is the skill that needs to be refined, will probably not benefit by submitting their growth plan in advance and by not having any discussion or changes made to it after consultation with the evaluator.

In the updated (1990) version of the guidelines, the purpose statement has been removed. The "how" of the Board's evaluation process has been lengthened but the "why" has been eradicated. Another separate, lengthy Board document (titled Growing Together) outlines the purposes; all the same, it is unusual that a detailed statement is not made on the evaluation document itself for teachers to read.

The first page of Growing Together reviews a new Board resolution that recognizes the need for a resource booklet to be developed to assist teachers and evaluators (April 18, 1990). Unfortunately, for the majority of my colleagues and myself, this document was never distributed as a tool to assist us in our respective roles in the evaluation process, even though I was called on to evaluate two of my own department members. The second page of the document is titled "The Main Purpose of Evaluation". What follows is a lengthy quote by Casper F. Paulson Jr. (source unknown) which poetically compares the change process to sailing a boat:

I would like to think that those of us committed to the educational enterprise have not surrendered to a mindless gravity, nor do we trust in a manifest destiny, nor are we driven like flotsam across the sea. I would like to think that we can "put our sheets" to the "winds of change"... for motive force. But how then do we set our course and pursue our destination? We must measure the wind, read the compass and use the sextant. We must chart and log our course. And sometimes in the dark of the night we must use celestial guides to steer toward our terrestrial objective. And that is what evaluation is all about.

I can only speculate as to why this nebulous nautical passage was chosen and titled as "The Main Purpose of Evaluation". Perhaps it is merely a dramatic and metaphoric introduction to the otherwise dry consideration of policies and procedures that follow in the handbook.
In the following section, the primary purpose of teacher evaluation is stated as a way to insure that the Board's “mission” will be achieved: "Each [deleted Board name] Public School is committed to educate every student in its care by providing the skills, knowledge, and values necessary for individuals to be capable and concerned participants in Canadian society". Four objectives are listed next related to effective teaching: to monitor and ensure effective delivery of programs and services to students; to maximize the effectiveness of teaching staff; to promote ongoing professional growth and development in educators; to obtain information for making educationally sound management decisions such as future staff development plans, the selection and development of candidates for promotion and, when necessary, the intensive assistance and documentation of those teachers whose teaching performance is unsatisfactory. The purpose has been updated to recognize that teacher evaluation also includes not only a pedagogical purpose in improving performance but a political and professional purpose in ranking teacher’s abilities with regards to possible promotion or possible documentation of poor performance. The Board acknowledges the implicit purpose of evaluation as maintaining the power structure and hierarchy found in this school system where evaluators would decide, from classroom observation, who will advance in the system. The formative purposes of evaluation, with its focus on improvement, and the summative purposes, with its focus on judgment, have been combined. While some researchers in the evaluation field (e.g., Darling-Hammond, 1983) advocate for the separation of these two motives, this school Board continues to fuse them.

Criteria for Making Judgments

The remainder of this Board document Growing Together deals with probationary and permanent contract legalities and rules regarding the distribution and submission of reports. It also includes the template for the written report that follows a classroom observation. On the back of the template are the revised criteria and standards for making judgments when preparing the written report (Appendix A).

These criteria are divided into those that should be included and those that might be included in the
teacher's written report in recognition that teaching is far too complex a task to examine all aspects at once (Cangeloni, 1991). It was reassuring to see a number of fundamental changes in the criteria from those used in the former version. Many of the descriptors listed as “observable indicators of effective teaching” that focused on the teacher's values and personality had been removed, such as "warmth" and "enthusiasm". Other changes included the removal of “responsiveness to individuals” and “absence of negative criticism”. “Classroom control” was rephrased into the more positive “management of learning environment”. “Use of student answers in the development of the lesson” had been replaced by the more neutral and encompassing “student involvement”. “Correction or remediation” was new to the list of indicators, as was “attention to different learning styles”. New theories of knowledge acquisition and evaluation as manifested by the inclusion of these new standards. Student evaluation was regarded as a new important focus in improving instruction with many separate subheadings to guide the assessment, e.g., diagnostic techniques, formative techniques, summative techniques, reporting to parents. In the previous guideline, “evaluation of student progress” was an optional criteria to be used in a discussion with the teacher.

A consideration of the additional criteria that supervisors could use in a written report in both the old and new evaluation guidelines demonstrates that the Board’s administrators had decided to include more specific out-of-classroom criteria in teacher assessment. The term “professionalism” was dropped from the more recent list. Also removed were the criteria “quality of program”, “learning environment”, “student assignments” and “realization of objectives”. “Interpersonal relationships” was modified to the more definitive “peer inter-relationships-colleagues, administrators, support staff”. New components include “supports school policies, supports school objectives, assumes duties outside the classroom, extracurricular involvement, knowledge of system philosophy, expectations, and program, evidence of growth, and, keeps abreast of new development-participation in professional association”. More than half (12 of 22) of the standards for making judgements are placed explicitly outside the realm of classroom teaching performance. I would argue that many of the classroom indicators are impossible to measure from a one time collection of data, notably, knowledge of current teaching/learning strategies”, “use of routine”, “reporting to parents’ and “knowledge of subject/program".
Number of Classroom Visitations

Even a minority of the criteria outlined above would necessitate more than just a single classroom observation. Traditional measurement perspectives surmise that increasing the number of observation occasions will increase consistency and reliability (Rowley & Simon, 1974; Stodolsky, 1990). Planning, evaluation and instructional methodology cannot be determined without considerable discussion with the teacher outside of the classroom. The actual written reports revealed an increase in the reported classroom visits and private conferences held with the teacher as part of the evaluation process. From the nine elementary reports that documented the number of visits, six identified that there had been two or more classroom visitations. Six reports stated that two private conferences had been held with the teacher, one report stated that only one conference had taken place and two reports stated that no conference had occurred. The trend in the elementary panel may be towards more one-on-one discussion with the teacher concerning his or her philosophy, methods and evaluation.

Each of the secondary written reports involved only one classroom visitation. All but one of the secondary reports claimed to have included two conferences with the teacher. I can only speculate as to the reasons that would explain the difference in elementary and secondary procedures. Elementary lessons are often only 40 to 45 minutes in length while a typical secondary class runs 75 minutes. Perhaps elementary administrators feel that more time is needed to observe their staff in their generalist role. Most elementary teachers teach a variety of subject areas and this fact may account for the administrator's desire to visit a class on more than one occasion. Another more pedestrian reason for the difference in time observed between elementary and secondary teachers is simply the reality that the number of staff in an elementary school is far smaller than a secondary staff; thus the administration may have more time and energy to devote to evaluation procedures. There also may be few immediate rewards for some secondary principals to improve their teachers' instruction.
Pre conferences

Pre conferences with the teacher to be evaluated are typically observed in this Board, perhaps due to both Union and Board policy that states that performance appraisal must include provisions for pre conferencing whenever classroom observation is used as a method for gathering data (OSSTF, 1993; Board Policies and Procedures, 1992). The teacher may present a detailed lesson plan to the administrator at that time. The administrator may ask questions to reveal additional information regarding other non-classroom criteria such as contribution to total school life and professional growth activities and should provide feedback and suggestions concerning the upcoming lesson that he or she will observe. In Lawton, Hickcox, Leithwood and Musella's (1988) study of performance appraisal in Ontario schools, only 36 percent of the teachers indicated that a pre-evaluation conference took place, with the average time allotted being 20 minutes.

Post Conferences

The tendency towards more conferencing with the teacher before and after the classroom observation could be perceived as a positive step; however, the practical realities of the process as I have experienced them, and researchers have noted, often reveal disappointing facts. Keith Acheson and Meredith Gail (1987) suggest that post evaluation feedback is extremely critical. It is an opportunity for the evaluator to show his or her objective data and observations to the teacher without comment. The teacher should then be encouraged to assess the data and comment on their strengths and weaknesses. Together, the teacher and supervisor can agree on alternative approaches for the future which the supervisor helps to facilitate. Although all of these 20 reports recorded that a post conference had occurred, in my own personal experience, and in the experience of the majority of teaching staff that I have had interaction with, the post conference is typically held as soon as the report is written. The purpose of the post conference becomes merely to provide the teacher an occasion to read over the report and to sign it, stating that he or she has received a copy; the purpose is not to discuss the contents or areas of growth, or a subsequent support plan. This experience is consistent with the Ontario studies of Lawton, Hickcox,
Leithwood and Musella (1988) who found that, "the incidence of planned follow-up activities is extremely low" (p. 5).

Problems in Credibility: Subverting the Process of the Classroom Visitation

Other disappointing trends emerge from my own experience in this school system, namely that four administrators did not stay to observe an entire lesson, two did not even show up for the scheduled lesson and one merely asked for a lesson plan on which to base his report because he did not intend to watch the lesson. Perhaps these administrators perceive this sabotaging of procedure as the ultimate in flattery to certain staff members. With a conspiratory wink, the administrator might inform a teacher that while a report will be written, there really is no need for him or her to actually be a presence in the classroom, since that teacher's competency is well known to him or her. These actions by administrators are in clear violation of a Board policy that outlines specific steps of the teacher evaluation process. Along these lines, Richard Stiggins (1988) argues that classroom observations have to be sufficient in number and duration to produce a representative sample of performance. Milbrey McLaughlin and Scott Pfeifer's (1988) study of four school districts in California reports other administrators shirking their evaluation duties:

I had only one observation (this year), but I never had a chance to sit down with my evaluator and look at what he wrote. This year he just caught me in the hail and said "I'm going to drop by and see you sometime this week." Then two weeks later he dropped into my class unannounced . . . several days later he stopped me in the hall and said "You had a great observation." To be honest, I felt somewhat brushed off. In fact, I was downright offended because when you evaluate someone that way, you're basically talking about the dignity and worth of the individual. (p. 126)

Teacher alienation is also reported in Ronald Doll's (1983) Supervision for Staff Development:

I asked my supervisor why he was visiting me. He said, "Because the school Board rules say I have to". "What do you do with your notes?" I asked (as thought I didn't know). "They go into your file, but don't worry, they're complimentary enough," he replied. I guess my file is more important than I am. Strange world, isn't it? (cited in Giles & Proudfoot, 1990, p. 259)

It is troubling to think that even in the most unusual and infrequent cases, administrators in any jurisdiction might circumvent board policy, ignoring their own professional duties as well as jeopardize any respect they may have developed with their staff. It is extremely difficult to define
the extent of this type of action even with the complete records of this Board, since only a date had to be included as proof that a post conference occurred. I venture that lack of time due to the multitude of principals' responsibilities, the extra effort required to complete a debriefing along with a lack of desire by perhaps both the teacher and administrator to lengthen the evaluation process may lead to various forms of subversion of intended Board policy. Perhaps this phenomenon is just another example of the complex relation between policy talk and actual institutional change.

Teacher's Written Feedback

Comments that a teacher is willing to express about the evaluation process may be found in the written feedback area of his or her report. Five blank lines are included on the front of official teaching reports (see Appendix A) to provide the teacher with the opportunity to make comments. Only four teachers from this sample made any statements. All four of the teacher's comments were to thank the administrator for a positive evaluation experience: "I have very much enjoyed the supportive administration, staff and community"; "I feel that the evaluation process was a positive one, and a beneficial learning experience for me"; "I appreciate the time and effort that Mr. R. put into our conferences and discussions"; Mr. K. is a very supportive and inspirational Head of Department. He made the evaluation a positive, learning experience". One teacher attached a separate sheet, summarizing and reflecting on the procedure:

The evaluation process used this year included a group review of the process and criteria, a personal interview, a classroom visitation and a review of the written report. This provided me with the personal assurance that many factors and facets of teaching were taken into account while the evaluation was done. I would also like to thank Mrs. S. and Mr. Q. for giving me the opportunity to seek new challenges and directions for personal and professional growth and having the confidence in me to support my desire for promotion.

None of the comments reflect a true consideration of the lesson observed, teaching strategies, reflection of the assessment or possible growth plans. What they do demonstrate in this case is the teacher's gratefulness in receiving a positive report, or, perhaps, an attempt to gain favour with the administrator especially since he or she is seeking a promotion.
Only one teacher had attached a separate sheet to her evaluation that added specific information she wanted to be submitted to counter some rather negative statements made about her teaching skills and professionalism:

I wish to add under INSTRUCTIONAL METHODOLOGY that the biography in question was jointly chosen . . . by the classroom teacher and myself . . . In the section CLASSROOM MANAGEMENT, I add that mention was made that the next step in the lesson NOT be stated . . . . I add to CONTRIBUTION TO SCHOOL LIFE that I also planned and conducted . . . . I include, in the AREAS FOR GROWTH, that Ms F. reminded me only once of the "total silence" regulation. I add that at times, as all teacher librarians have found, it is essential to work at a desk where a computer is located, in order to assist students . . . . It is unfortunate that Ms F. did not feel that being notified on . . . was not enough lead time to reorganize the library schedule.

This teacher felt misjudged in many of the criteria areas, as dramatically noted by her use of block caps in the text. I am left to wonder why the administrator and department head who have been critical of her in these areas, did not discuss and change the comments in the official report, and instead, left them as a separate submission by the teacher in question.

Length of Written Reports

With such a long list of mandatory criteria in addition to a substantial supplementary list for the administrator to use in assessing the teacher's performance, I expected these written reports to be detailed and lengthy. The average length of these reports was two pages that totaled 300 words. The longest report belonged to an elementary teacher at five pages. The shortest were two secondary reports, both that were less than one page totaling 110 words. Perhaps the evaluators express satisfaction with teacher performance by writing a concise statement. Some evaluators may feel that a couple of paragraphs are sufficient to get the teacher's mandatory evaluation completed and back to the Board office before a deadline. Still others take a significant amount of time, care and effort in completing a detailed observation and assessment of the staff member based on a variety of sources.

The length of the report was not necessarily an indicator of relevant detail, though. Many of the longer reports were written in a vague style. These statements, as examples, were common; "School A's staff and students have benefited in many ways from Ms. W.'s contributions to the
school. She has demonstrated efficiency in all the aspects of department administration. Ms. W. provides an outstanding example of professional behaviour for members of her department”. Lists were also very popular in the written report, frequently lacking adequate description of the items indicated: “The students are evaluated using the following methods, summative tests, formative tests, formative teacher evaluation of students in the group, co-operative learning projects, student self-evaluation”. In four of the reports I found what I thought was excessive detail that focused on describing the physical qualities of the classroom while ignoring more important facets of the teacher’s performance:

Many display materials relating to her program and themes are evident – Cinderella stories and pictures, an octopus student duties board, bear poems, graphs, patterns etc. The room also contains many centres such as computer, listening, reading, games, puzzles, building, cut and paste. It is a very colourful and inviting room.

While a description of a conducive learning environment is important, in this case, it was far outbalanced by the remaining two paragraphs of the report that summarized a Halloween story that was read to the students.

General Assessment Ratings

In the updated policy on evaluation, guidelines for granting a "general assessment" to teachers were changed to differentiate between probationary and permanent teachers:

General assessment ratings must be shown for all probationary teachers and for permanent teachers whose work is unsatisfactory. Others are optional.

For Probationary teachers only:
Satisfactory: Meets Board standards
Unsatisfactory: Fails to meet Board standards
Note: A probationary teacher requires a general assessment of "satisfactory" in order to qualify for a permanent contract.

For Permanent Teachers Only:
Excellent: Exceeds Board standards on a regular basis.
Very Good: Meets Board standards on a regular basis.
Satisfactory: Meets most of the Board standards on a regular basis.
Unsatisfactory: Fails to meet Board standards.
Probationary teachers can only rate a "pass" or "fail" in their performance in a determination of whether they would be hired. The other notable change is that permanent teachers were now to be judged by Board standards (e.g., "Meets most of the Board standards on a regular basis", 1990 policy) instead of the more narrow definition of teaching performance (e.g., "A teacher whose performance is without any major weakness", 1980 policy). The Board wanted to include all aspects of its policies and procedures in a judgment of what makes an effective teacher. No documents define the Board’s standards.

A teacher, as an employee of this particular Board, must adhere and promote all aspects of the "system philosophy, expectations, and program". One report in particular stated that the teacher was “a positive team member who supports school, area and Board initiatives”. Fifteen of the reports included the term “team player” or “team builder” as positive comments.

Perhaps as a result of this stress on Board standards as the overall indicator of a teacher’s success, teachers that do not share their employer's educational objectives might find themselves with a negative evaluation in their file at the Board office. While most educators share a majority of educational goals (e.g., the students' well being, the fostering of intellectual and creative development), there may be significant disagreement as what are the means to achieve these ends. If a board, for instance, decides that destreaming is the best way to teach students and a teacher disagrees, the teacher, according to the new evaluation guidelines, could now be considered as non-supportive of that board. This is one of the most striking changes of the Board's evaluation policy within only ten years time. Its language and intent is borrowed from corporate models where employees are expected to support management's goals, or lose their job for not being a "team player". This change indicates a return to the more traditional, summative evaluation of teaching professionals.

Despite the clear specifications on who can be awarded which type of overall assessment, the reports divulged the following discrepancies. One elementary probationary teacher was rated as "excellent" (this rating was not to be used for probationary teachers), one was rated as "satisfactory" and one permanent teacher was rated as "excellent". The remaining seven permanent elementary teachers were given no rating at all. Two secondary permanent teachers were rated as
"excellent". The remaining eight received no overall assessment rating. None of these teachers had an unsatisfactory rating, even though one report in particular was very negative regarding the teacher's lesson and overall professional behaviour (e.g., the lesson's "suitability for English as a second language students was questionable" and "students were confused" and "there are concerns about Ms. Y's interpretation of her job description and her compliance with department policies"). Administrators, for unknown reasons, chose to inconsistently apply or ignore the Board's own policy in regards to assessment ratings.

Criteria used in Teacher Assessment

According to this Board's procedures and the Ontario Secondary School Federation's guidelines, teachers are to be made explicitly aware of the criteria by which they are judged before the classroom visitation. Richard Stiggins' (1988) studies of teacher evaluation in Ontario found that one of the main factors in its effectiveness is if the evaluation is related to relevant predetermined performance criteria, criteria that are uniquely relevant to the teacher. Lawton, Hickcox, Leithwood and Musella (1982) determined in their survey of Ontario schools that 71 percent of teachers indicated that the purpose, but not the criteria, of their last evaluation had been clearly communicated. The study cautions that even this positive result is "somewhat ambiguous given that not much time is usually spent in preparation for the evaluation of teachers" (p. 21).

While the mandatory criteria listed in the Teacher Evaluation Guideline aim at the stated purpose of improving instruction, the optional criteria focus on the professional growth of the teacher. An analysis of the administrators' comments revealed if criteria headings were used, which mandatory criteria would be the focus of the reports and what proportion of the written reports were devoted to discussing mandatory and optional criteria. If thoughtful Board philosophies and educational purposes dictate the criteria on which to base an assessment, surely the administrators would follow these standards when making their judgments. An accompanying evaluation binder, Growing Together, was published in 1990 as a tool to assist administrators in this goal of standardized assessment criteria.

The ten elementary reports had interesting results reflected in the use of criteria. A total of five of
the reports used no criteria headings in the written report. When reading the report, I would have to read between the lines to identify criteria that were implicit in an anecdotal review of a lesson. The other five reports used criteria that were not listed in the policy document, such as "lesson presentation", "quality of program" (versus the policy's statement of "knowledge of program"), "classroom atmosphere", "instructional leadership", "management functions" and "professional relationships". While these terms might be close in intent to the stated policy, that administrators choose their own terminology when reporting criteria suggests that even the best board policy cannot ensure its implementation at the school level. Only three of the reports mentioned student evaluation policies. I can only speculate how instructional practices can be improved when there is no mention of why and how the students' learning is assessed.

Three secondary reports had no guiding criteria headings and were a synopsis of one particular lesson. The other seven evaluators were careful to select from the Board's mandated list. The two most frequently used criteria were "classroom performance" (an obvious choice considering the mandatory visitation) and the other, perhaps more surprising, criterion that was used in every single report, was "contribution to total school life".

It appears from this sample that what is important to the evaluator as administrator is what the teacher does outside of the classroom. Whether the teacher is a coach, chaperone, attendance monitor or chairperson is considered a crucial area of their assessment. The following examples illustrate the importance of extracurricular performance in a teacher assessment:

Ms. E. is presently working with our Folk Dancers and is becoming more involved outside of the classroom.

She has been an active contributor in a variety of ways. Her contributions in these areas have been greatly appreciated.

Ms. J. often spends her lunch hour at school coaching teams and clubs.

She has coordinated a number of school initiatives on special occasions (i.e. Education Week, Heritage Week, electives, assemblies, year end outings). She has coached bantam boys, girls and coed volleyball as well as running long jump for track and field. Mrs. R. assists with the newsletter in writing articles for it as well as being one of the main proof readers before printing.
Being an excellent classroom teacher does not appear to be enough in the eyes of these administrators and Board officials. Ideally, one must contribute to total school life. The reality of the school system today is that it is cannot be run without the unpaid administrative assistance and voluntary extracurricular work of its teachers. Comments in the reports reflect the administrator's recognition of this fact: "We are appreciative of the extra hours invested". Frequent acknowledgment is made of the "countless hours" that staff members put into "non teaching" activities.

Besides the other non mandatory criterion that was used in 17 of the reports was "Professional Activities and Professional Growth". This would seem to be a positive trend since the stated purpose of the teacher evaluation is to encourage professional growth. In service training, workshops, professional reading and conferences were some of the areas in which the teacher's past efforts were acknowledged: "Mr. G. has attended in-services in such areas as Later Literacy and Reading Recovery", and "Ms. S. has also attended several Professional Development programs relating to the Transition Years . . . and then shared the information with us".

Recording Future Growth Plans

A problem that exists in many of the reports involves the future growth plans of the teacher. Many vague statements were made such as: "He sees leadership skills as an area he would like to develop"; "Continue with your efforts on assessment and evaluation. . ."; and " . . . improve the classroom music program". This absence of more specific suggestions and guidelines for the teacher's professional development reflect poorly on the evaluators and their own familiarity with appropriate professional opportunities for their own staffs. The sketchiness of these type of comments also confirms research that finds that "formative evaluation is one of the most poorly developed features of the entire range of supervisory responsibilities" (Daresh & Playko, 1995, p. 290). This lack of formative evaluation may be due the expectation of an open discussion of deficiencies as well as strengths in the cause of improved performance, that is far too large an investment in time and interpersonal skills (Daresh & Playko, 1995).
One vice principal was candid in his comment that the teacher should attend upcoming conferences and workshops "that are financially feasible". Six of the evaluators had computer literacy as an area of growth for the teacher, making technology the most frequently mentioned area of growth.

Should individual teachers be held completely accountable for their own professional growth? Recent research by Kremer-Hayon (1993) has made a viable connection between a school's climate and its powerful effect on the professional growth of individual teachers. Whether the school is supportive or restrictive and closed affected not only collegial relations (Little, 1982, cited in Kremer-Haron, 1993) but how effectively change was implemented and the degree of teacher development (Fullan, 1982). In a climate where principals are just as responsible for teacher and school improvement as the classroom teacher, I think that teacher evaluations would be fundamentally different than the selection of written accounts that I examined. Perhaps the administrator who suggested "financially feasible" workshops might offer encouragement to the teacher by assisting with the school's allocated professional development funds.

In two of the reports, the evaluator suggested as a growth plan that the teacher would be good candidate for promotion. In the 20 reports that I read, it became very clear which teachers were candidates for promotion and which were not. In the opening paragraph of one written report, it was stated that Mrs. X. "would be an excellent candidate for the position of vice principal and would have my full support should she decide to do so". Yet another stated: "Mr. G. expressed an interest in applying for promotion to chairperson. We will ensure that he is made aware of possible leadership opportunities as they become available". No written advice was given regarding the steps needed to attain this goal of promotion.

I have spoken to many teachers who felt obliged or were told directly to include a personal "growth plan" with the report. This "growth plan", instead of being an opportunity for the teacher to ponder on his or her effectiveness as a teacher and to articulate future professional development, is typically only a set of blanket statements listing teaching workshops that they might attend in the future. This is due, in part, to deadlines on the submission of teacher reports that hinder the teacher from writing a truly reflective piece. The official evaluation form has only a limited space for the growth plan. It is not part of the Board policy to hold teachers accountable for any of these
statements (interview with supervisory officer, 1999). It is left up to will of the teacher to pursue any stated goals on an evaluation report.

Problematic Observations

In what may be an understatement, Popham (1988) notes the problem of maintaining classroom observation as the primary means of teacher assessment: "Even though classroom observations are almost universally regarded as the sine qua non of a complete teacher evaluation system, they are not without serious problems" (cited in Cangeloni, 1991, p. 47).

One of the basic problems, perpetuated by evaluators since the 1870s, is the continual use of personality traits as valid criteria for what makes an effective teacher. In all of the elementary reports, the administrator had made a minimum of five to six of these comments. Four of the reports had over a dozen of these types of statements and descriptors (e.g., pleasant, funny, caring): "Mr. D. is a real breath of fresh air on our staff"; "Her enthusiastic personality and originality offer a great role model for her pupils"; "Her caring attitude, flexibility and good humour giver her pupils a great year"; "She is friendly and supportive"; "She is very empathetic with students", and "Her calm, motivational and supportive nature is instilled in her students".

The secondary evaluators, on the other hand, were less inclined to articulate these types of observations. Three of the ten reports made no reference to personality while the remainder made only two or three comments: "Mr. E. demonstrates a warm and open-minded manner in his dealings with students" and "Miss. Y. is very caring with many of the students in her class." Perhaps the evaluators felt that due to the tender age of the children in their care, elementary professionals themselves must have these positive character attributes in order for the children to learn. In secondary classrooms, with its focus on subject-based expertise and preparing young adults for the world of work or post secondary training, personality appears to be considered a less worthy criterion for student learning. Regardless of the reason, what evaluators are perhaps unaware of is the enlarging body of research that proves that "teaching styles and personality characteristics appear to vary as much among teachers judged as highly competent by students, supervisors, and peers as they vary for teachers in general" (Allen, Davidson, Hering & Jesunathadas, 1984, cited in Cangelosi, 1991, p. 10).
Gender Differences

The difference between the descriptions made of male and female teachers were compelling. The seven male teachers were described as thorough, detailed and precise, caring, respectful, approachable, leaders, competent, professional, resourceful, proactive, well organized, having a good sense of humour and self-sufficient and hard working. Curiously, one evaluator felt it necessary to note one male physical education teacher's "exceptional athletic ability".

Women, on the other hand, were frequently described as caring, pleasant, bubbly, calm, friendly, supportive, empathetic, controlled, sensitive, open, poised, diligent, approachable and enthusiastic, endowed with a high energy level and able to use her voice effectively. One teacher was characterized as "a role model, especially for female students". Another elementary teacher’s official assessment report records, "her positive and bright outlook that has helped cheer up many a staff member". It becomes evident that some evaluations are marked by stereotypical attitudes of gender roles both in and out of the classroom. A woman's personality in the classroom was commented on in much more detail than her male counterparts. As Merill (1983) and Borisoff & Victor (1997) noted, gender and gender expectations may partially determine how supervisors interact with those whom they supervise. Research tells us that the sex of participants affects what is communicated and how it is communicated.

Carol Shakeshaft’s (1989) considerable research proposes gender as an explanatory variable in theories and practice in supervision. Since men and women communicate differently and focus on different information, this variable may result in biased assessment procedures: "It may be the case that in a supervisory conference in which a principal is discussing an instructional issue with the teacher, the women participant is listening for the feeling and the man for the facts" (p. 143). One report that I read did comment that a female teacher had, “maintained a highly professional, receptive attitude towards all suggestions offered” during a post conference. It remains unclear what else the administrator might have expected from her.
Shakeshaft also speculates that the different values that male and females carry into the job may lead to women focusing on instructional and child centred issues and men choose administrative issues to discuss with the principal. Because I could not accurately identify the gender of the evaluator in each case and I have no record of the oral communication that occurred during the assessment procedure, I cannot comment on her assertion that male teachers exhibit more hostility in dealing with female administrators than do female teachers.

Shakeshaft’s (1987) analysis of written reports finds women who are initially evaluated less favourably than equally competent men and that males evaluate females more harshly than females evaluate females. She found women receive more neutral forms of feedback while men receive both more positive and negative responses. The small sample that I reviewed contradicts these findings. Women had more positive statements made about their personalities, classrooms and teaching styles. A female principal (the report indicated a first name, unlike others in the sample) gave the most critical report to a woman. The comments made about the men were more neutral. To illustrate the difference, the following excerpts taken from one male and one female’s report:

Mr. R. is presently a teacher at School B. Mr. R. has acquired an exclusive and comprehensive breadth experience in teaching young adolescents in a number of senior schools within the city. Mr. R. believes that as a teacher, his major function in the school is to work as a member of a team to help socialize young people to acquire the skills, knowledge and attitudes that will allow the student to mature as a responsible, healthy and happy individual within our community. He recognizes that the task is indeed very challenging in our modern society, yet he is prepared to direct his energies and resources to fulfill our mandate in teaching.

Mrs. A. continues the growth she displayed in the Fall. She has a great rapport with the pupils, displaying her warm, caring attitude. She is still well organized spending considerable time to achieve her goals. Her enthusiasm is infectious and is very evident at all times. She is quite flexible which is really the mark of a good teacher couples with a sense of humour that is so essential. Her classroom continues to be a pleasant, attractive learning environment.

While chosen to highlight differences, this type of commentary, describing Mr. R. as a teacher with considerable experience, articulated teaching philosophies and goals and Mrs. A. as a teacher whose personality traits are intricately connected with her success as a teacher, are discerned in the same degree in the other reports. Perhaps a larger sample as well as an indication of the gender of those writing the report would show different patterns.
The traditional view of women teachers as nurturing role models and male teachers as those who impart intellectual wisdom has not changed as much as one would hope in over two centuries. Another difference found in the reports of male and female teachers was that in three of the female elementary reports, the evaluator included in the assessment of the learning environment the actual physical surroundings of the classroom: they were described as "warm and comfortable", "colourful and bright" with details about artwork, accessories and bulletin board displays. I would conclude that gender differences were evident in the descriptors and focus of the reports.

Observing the Learning Environment

Aside from descriptions of the physical space as well as the characteristics of the teacher, the majority of evaluators provided exacting detail of the teachers' actions throughout the lesson. Infrequently, there were statements made about the members of the class: "Students worked enthusiastically and cooperated"; and "Students were focused"; and "Some students were confused". These types of comments are problematic. Far from being objective accounts of what the evaluators witnessed, they are judgments of the teachers' effectiveness. But do the evaluators actually know if the students are learning or if they understand what is being taught by watching them? One administrator wrote, "Everyone worked to the best of their ability", a far fetched, unsubstantiated statement lacking hard data.

Further comments on the reports are comparable in content to those found in the earlier political liberalist tradition of the mid 19th century: "Students speak openly and politely"; "They worked assiduously, completely involved in their projects, totally focussed and enthusiastic about finishing their designs"; and "Students are seated in individual rows". By commenting on student behaviour (and seating arrangements), and ignoring performance assessment of any kind, the evaluator cannot make many valid judgements concerning the learning that occurred, only that students were under control, well mannered and conforming to the expectations of a traditional disciplined classroom.

Evaluators need to heed the warnings of Cangeloni (1991) concerning writing reports based on what is witnessed in the classroom:
The classroom with its social structure and activities is an incredibly complex environment. Far too much goes on in a classroom, much of which is influenced by unseen, unstable factors (e.g., whether or not a student argued with a parent the night before), for classroom observations to be the "direct" data source relevant to teaching competence, teaching performance or student outcomes that they are sometimes naively thought to be. Unless extensive efforts are undertaken to educate and train observers to use well-designed observational instruments, classroom observations will continue to be dominated by malpractice that produces invalid results. (p. 47)

Many practitioners share Cangeloni's (1991) concerns when they realize the importance that one classroom observation plays in the overall assessment of a teacher. A Friday last period evaluation would undoubtedly be different from one on a Monday morning. Observing a class in September is quite different that observing it December or June. While most principals, as former teachers, should be sensitive to the invisible factors at play in the classroom, what of department heads who are called on to evaluate their peers in their own timetabled preparation period? In my experience with this particular Board, I never had the opportunity to be formally or informally trained in assessment procedures, design or implementation even though I was called upon to evaluate two of my own department members.

Problems in the Language of Reporting

I had initially anticipated interesting summaries of teachers in action when reading these documents. Fuzzy criteria, imprecise descriptions and even cliched statements were prevalent, however. Cangeloni (1991) also observed this phenomenon: "Most classroom observations either provide data for summative evaluations that are too vague to be defended or formative evaluations that are not specific enough to influence teaching performance meaningfully" (p. 231). The selected reports I read all contained frequent generalities.

The following are a small sample of these statements that gloss over particular aspects of teacher effectiveness: "She is a very hard worker"; "She has many innovative ideas"; "Her written reports are always well done"; "Mrs. P. is very comfortable in the classroom with the curriculum and the students"; and "Mrs. C.'s student evaluation is of high quality". No evidence or even a single substantial example is provided to back up these statements. Lawton, Hickcox, Leithwood and
Musella (1988) had comparable results in their study: first, a tendency to use vague and ambiguous criteria and second, criteria that often had no real relationship with effective teaching or learning.

In the 20 evaluations of this study, statements were often made which left the reader confused in regards to the evaluator's exact meaning, but, perhaps, that was his or her intention: "He perceives himself as a team player and a good listener"; "Her sincere desire to have a positive impact on children has emanated from all facets of the process"; "He generally completes tasks in a caring manner, with few errors". A principal once told me that in order to avoid confrontation with a difficult teacher, many evaluators write in this “read between the lines” style, in the hope that the next administrator can pick up the nuances of what has been written. A teacher that “perceives” himself to be a team player and good listener is not the same as the teacher who is a team player and good listener.

In six separate documents, the evaluators commented on the teacher's philosophy e.g., "Mr. X. believes that students should develop into life long learners . . . ." and "She continues to be concerned about goals and objectives". While it is commendable that the teachers could articulate their educational philosophies, there was no solid evidence stated by the evaluator concerning specific professional endeavors that would demonstrate that the teachers were in fact putting their ideas into practice.

The last major problem in the language of the reports was the use of absolute descriptors. Teachers, after the one period observation, were described as always prepared, always calm and constantly adapting curriculum. Four other faultless teachers were depicted: “Ms. D. never has discipline problems”; “. . . she is constantly searching for new ways to help students increase their knowledge and skills . . . .”; “She is continually seeking innovative instructional practices to improve delivery of the Art program”, and “He is constantly searching for appropriate texts and other resources”. In an attempt to be positive, the evaluator’s choice of language has made the comments ineffective.
Positive and Negative Statements

What became increasingly apparent as I read through each report was the congratulatory tone. All but one of the reports read in parts like a thank you note to the teacher. Harold Guthrie and Donald Willower (1973) observed this same phenomenon in their study of Pennsylvanian teaching reports and aptly named it "the ceremonial congratulation". The summaries were very positive and laudatory. These are typical examples: "We suggest he clone himself"; "Ms. A. is a real breath of fresh air on our staff"; "Hopefully, most of our new people will be like Mrs. B."; "Mrs. E. has been a fine asset to our staff and I wish her continued success"; "Ms. D. is an outstanding teacher and a valued member of School D's staff"; and "I have consistently been impressed by her background, knowledge, skill and sound judgment". The most frequently repeated statement in 12 of the reports was that a teacher was "an asset" to the school. While the principal should be applauded for making such encouraging comments to their staff, there is a lack of critical statements. It seems that the administrator has an entirely different purpose in teacher evaluation than merely improving instruction or encouraging professional growth. The principal, as an instructional leader, is keenly aware of the impact that his or her statements might make on individual teachers. If the principal is considered supportive and appreciative, teachers may in turn be more of an advocate of his or her administration.

The only purely negative statements were found in one report of a teacher-librarian: "Ms. Y. doesn't enforce the total silence rule"; "...sometimes there has been a lack of communication with the Head which has caused concerns..."; "It is recommended that Ms. Y. communicate more openly and frequently with her Department Head". In this single case, the principal stated two problems with the teacher's effectiveness and made a specific suggestion on how she could improve her performance. Referred to previously, this teacher's four paragraph response to the specific criticisms in her submitted report stated that she had only been informed once of a total silence rule and that she felt that she communicated regularly with her department head. Perhaps this one negative evaluation illustrates one of the barriers that evaluators face. If the evaluator is critical in a written report, a teacher has every right to refute the observations and suggestions made, unless they are of such a serious nature as to be considered gross incompetence and thus
they become subject to more serious Board policies and procedures.

Positive Observations

I must admit that I was surprised by the extent and variety of problems evident in the written reports. I expected the majority to be detailed, professional and helpful accounts of a teacher's abilities accompanied by specific attainable suggestions for improvement with a reasonable plan for current professional growth. What I read was disappointing in these regards; however, there were some positive exceptions to note. The increase in the number of classroom visitations points to many administrators recognizing the complexity of the teacher assessment task and being willing to place an increasing commitment of time into these assessments. Another positive trend was that two elementary school principals referred to a teacher's habit of making relevant parental contact. This is an area of professionalism that I feel should be assessed.

I could characterize six of the reports as exemplary since the evaluators backed up a majority of the statements they made concerning the teacher's performance with detailed references to the lesson that they observed. Two or three of the evaluators used observable criteria solely for making an assessment and avoided vague and questionable statements about teacher personality. These findings reflect fundamental differences among principals in observational skill and ability to judge teacher performance (Medley & Coker, 1987).

As mentioned previously, four elementary teachers wrote positive comments on the front of the teacher report stressing how the process had been beneficial for them. Only through a focus interview could it be determined whether these teachers actually felt professional growth had occurred or that they were merely reciprocating the "thank you" phenomenon illustrated in the report.

I was left to wonder how useful these written reports were to the teacher in promoting awareness and enabling professional growth if there were so many flaws and differences in criteria, style and focus. Performance appraisals are intended to yield useful information to the person being appraised (Musella, 1988).

Obviously, whether the information is used to assist the appraisee to improve
The reports seemed more of finality than part of an important process. Echoing McLuhan's famous insight about the medium being the message, Musella (1988) observed that choosing the form of communication relates to its purpose. A written report that is signed, duplicated and placed in a teacher's official file has a purely summative role. The written report's purpose can be to serve as an immediate reward and or punishment for the teacher. If this Board wants to facilitate goal-directed change instead of just record-keeping, significant education and direction is needed for the evaluators in addition to increased participation by the teachers. A sound appraisal policy in print is not a guarantee that the information collected and utilized meets the criteria of reliability, validity and usefulness (Musella, 1988).

Additional Areas of Study

There are countless potential areas for further consideration that were not within the scope of this paper that would add to the understanding of teacher evaluation. An examination of any trends towards punitive teacher evaluation in regards to non-adherence with board philosophies, expectations and policies could be analyzed. In order to gain a clearer picture on the entire process of assessment data, written and verbal exchanges of an evaluator and evaluatee through the entire assessment process could be documented. A comparison of teachers' attitudes towards formal evaluation is needed. An analysis of the training techniques in formal assessment practices for administrators may reveal interesting trends. Videotaping several classroom observations and comparing them to the resulting written report would highlight many problems in subjective written accounts. Examining a chronological sequence of teacher reports over an entire career to note patterns in the comments and the type of growth plans, for instance, may serve as one way to ascertain the validity of teacher assessment as a helpful process to teachers. A revealing comparison of written teacher reports from unionized and non-unionized staff may divulge notable differences. Finally, a consideration of many factors that might effect teacher evaluation policies that exist outside of the school, such as the relative strength of teaching unions, the state of the
economy and the actions taken by a Ministry of Education may reveal more about the intent of the process as pedagogical or political.

Legislated Changes that Reflect Educational Paradigms in Ontario

As examined previously in this paper, the progressive ideas and practices of the 1960s and 1970s were abandoned by some ministries and boards of education due to a perceived failure in the public and political arenas; they were replaced by conservative measures such as a reduction in optional courses in secondary schools, more standardized testing and a return to "the basics". A 1993 public opinion poll by the Angus Reid Group found that Canadians overwhelmingly supported a "back to basics" approach to teaching the fundamentals of literacy and numeracy, skills training and technology education (cited in Dunning, 1997). New emphasis on standarized outcomes and core subjects are reflected in the Ontario Ministry of Education's new curriculum from kindergarten to grade nine. The Conservative government, led by Mike Harris and then Minister of Education Dave Johnson, promote "rigorous core curriculum", "high standards", "school to work programs" and "increased time on math, science, English, and technology". Ron Manzer (1994) was correct in his assessment of the 1990s as shifting away from the ethical liberal ideology that guided policies of the late 1960s and early 1970s towards technological liberalism in a quest for educational excellence.

The demands for educational accountability seem to point to the continuation of teacher evaluation as a way to maintain minimum competency standards (Stodolsky; 1984, Haefele; 1992; Shinkfield, 1994). Both the Conservative Ministries of Education in Canada of the late 1990s, notably in Ontario and Alberta, in an effort to demonstrate fiscal responsibility, are questioning teacher professionalism, promoting a back-to-basics curriculum and implementing new province wide testing. A charter dated April 19, 1999 from the Premier's Office outlined this return to conservatism. This, The Charter of Education Rights and Responsibilities is a traditional treatise on education for citizenship that might have been written in educational policies for turn of the last century Canadian schoolhouses. The text outlines the responsibilities of students, teachers and parents in the education system. The charter further highlights the belief that there was a time in the past when there was more discipline in the classroom:
There was a time when respect, responsibility and discipline were more firmly rooted in society than they are today. These principles must be restored in our communities, and the best place to start restoring them is with Ontario’s youth. While these values are best taught at home, government can help by using the education system to teach respect for oneself, respect for the rights of others, and respect for the responsibilities of citizenship. (cited in a Canadian News Release, April 20, 1999)

As students in Ontario become more “disciplined”, the results of their learning and the corresponding performance of the teacher and the school will be tested in core subjects in every grade.

Standardized tests are one of the most accurate and reliable ways to measure student achievements and identify the areas where a better job of teaching needs to be done. That's why, despite opposition, the Harris government introduced standardized, province-wide tests in the core subjects of math, language and science in grades 3 and 6, with grade 9 to follow. Now that the test results are public, parents have a much clearer picture of whether their school is meeting the needs of their children. Being able to compare one school to another is just the first step. Parents should also know what the lowest-performing schools are doing about their shortcomings. (cited in a Canadian News Release, April 20, 1999)

Ontario schools are legislated to return to the 1920s and 1930s focus on standardized measurement as a way to categorize, stream students and become generally more accountable. Ostensibly, comparisons between schools will follow the competitive business model. The charter also states that students will not longer be promoted to the next grade level for any other reason than measurable academic achievement. These policies refute educational progressivism. The government also plans to begin the testing of teachers to gain standardized summative evaluation of practitioners:

The quality of a child's teacher can make or break that child's education. We have excellent teachers in Ontario but the world is changing rapidly and we've got to make sure all teachers are keeping up. They must have the up-to-date skills, training and knowledge to put our students at the top. It's common sense to make sure that Ontario's teachers are the best-qualified and skilled professionals. Working with the College of Teachers, the government will require all Ontario teachers to participate in a testing program to stay up to date, and require all teachers to take and pass recertification examinations every three to five years.

Radio ads were more specific in adding that teachers who do not pass the standardized test must attend summer workshops in preparation for a retest. Phyllis Benedict, the President of the Elementary Teachers' Federation of Ontario, seemed surprised by the Ministry’s plans, as seen in
her press release announcement:

The College, in a broadly consultative process, has been examining all aspects of teacher training and qualification. The announcement today, by the Premier and the Minister, was made without any consultation with teachers, with their federations or, apparently, with the College.

It will be interesting to examine how relations between the College, teachers' unions and the Ministry change in regards to the introduction and implementation of formal teacher tests. The liberal tendencies of Canadian education, as professed by Manzer (1994) are being replaced by more conservative policies. Teacher assessment may no longer require a classroom visitation but will also include a set of written tests.

While an increasing number educational theorists advocate for a shift in supervision away from traditional roles based on authority and inspection towards a philosophy that stresses employee participation in setting professional goals, this philosophy was not found to be in practice in the Board that I studied and will be difficult to implement soon in any Ontario public school boards. I predict that many of these political changes in the education system, as outlined below, might make it increasingly difficult for supervisors to pursue a collaborative growth model.

New legislation stipulates more of a managerial role for school administrators and moves them away from their progressive roles as master practitioners and curriculum leaders. Bill 160, passed in 1998, withdrew members of administration from teachers' unions in Ontario. Their duties remain with budgeting and maintaining curriculum standards and accountability. A principal or vice principal who is no longer subject to the values, rules and policies of their peers' union will perhaps feel more inclined to use summative assessment and other methods of coercive power on their workers. Many principals are finding it difficult to now facilitate change in a school system in which they are now defined as management. I know of a minimum of 5 vice principals who considered resigning their position to remain as part of a union. Concerns remain as to how they will manage the Ministry's demands for mandated teacher testing and still facilitate professional growth.

The small sample of reports from this Board illustrated a curious hybrid of the historical
educational paradigms from previous eras. The classroom observation, void of any true focus on professional development, smacks of the inspectorate era of Ontario education over 120 years ago. The use of general assessment ratings and the problematic criteria used in teacher evaluation policies were reminiscent of the scientific and technological thrust in education in the 1920s and 1930s. The stated purposes of this Board's evaluation policy combined the language of business management of the 1990s (e.g., "... to maximize the effectiveness of teaching staff") with humanist ideals of the 1960s (e.g., "Today's educator should be a self-assured and caring person with high self-esteem"). While the language declaring the purposes of teacher evaluation has become increasingly humanist-oriented and developmental, the actual practices and policies have yet to follow and still remain largely coercive and bureaucratic in practice. There needs to be more attention to revision and updating the newest theories in appraisal (Lawton, 1988).

Discussion/Implications

If formal evaluations must remain for all teachers, for political or pedagogical reasons, change in this Board and many others should be considered in the training of those who are the evaluators in teacher assessment. Every administrator should be made aware of the inherent problems in teacher evaluation that I have detailed from the sample reports and given explicit guidelines on how to avoid these problems. Some of the areas that deserve considerable training, reflection and practice should include avoiding generalizations, preventing a mere congratulatory tone, reporting in an unbiased manner on male and female teachers, making unsubstantiated claims, outweighing extracurricular contributions as compared to classroom performance and using criteria effectively. There needs to be some way to assess if the supervisory program actually results in improved performance. And most importantly for implementation, there must to be adequate time and financial resources allocated to support these changes.

The Board also should have training to reconsider chosen criteria. Often the criteria were vague, immeasurable and did not correspond to their stated purposes of teacher evaluation. In the crucial process of defining the criteria for effective learning and teaching, David Pratt's (1987) examination of the results of statewide testing in the United States made it possible to examine those schools that were at the high and low extreme in terms of student learning. The
characteristics that were typical of effective instructional settings include such factors as: a positive, orderly atmosphere; a clear academic focus; collegial planning of curriculum; active student involvement in learning; high student time-on-task; and appropriate difficulty of instruction. Policy makers and administrators would be well advised to use extensive educational research such as Pratt's in redefining clear, valid and reliable criteria for teacher assessment while shunning questionable criteria such as extra curricular involvement and personality traits. This Board also needs to rework in procedures in order to develop a balance between teacher practical autonomy and collaboration through the assessment process.

As recently as 1970, only 37 percent of Canadian elementary and secondary teacher held university degrees. By 1990, that proportion had risen to 80 percent (Canadian Education Statistics Council cited in Dunning, 1997, p. 43). With this important change in the qualifications of teachers as one of several important factors to consider, many educators and superintendents that I have spoken with think that the focus of formal evaluation should only be on the incompetent few and novice teachers. Researchers like Erlandson and Bifano (1987), Shinkfield (1994) and Airasian and Arlen (1993) agree. Teachers who receive poor reports should be prescribed assistance that reflects their needs (Bridges, 1986). I reiterate my belief that summative assessment, through teacher self reflection, peer review, portfolio submission, peer mentoring, student feedback and specific and realistic growth plans that are supported by an administrator would provide a valuable experience for a teacher. By exploring teaching practice both privately and along with peer support, real change and learning is more likely to occur for the teacher.

In sum, since the early 1870s, teacher evaluation has often been perceived by classroom teachers as being in conflict with several goals central to the teaching profession, namely, increased professional status and increased autonomy in classroom practice. Teachers have often questioned the purposes, techniques and results of having their effectiveness assessed. The correspondence of administrators through more than a century of educational change has also been critical of the enormous time commitment required to perform a task of questionable validity (Haefele, 1992). Teacher and administrators alike question the effectiveness of evaluating competent teachers year after year (Hickcox, 1989). While definitions of good teaching have only changed modestly since the 1870s (from disciplinarian to patriot to facilitator), the purposes and processes of evaluation
have not radically altered (Varga, 1991). Classroom observations, held in isolation from other possible criteria, have never been a reliable basis for evaluation, since no clear relationship exists between the learning process and teaching behaviours (Hickcox, 1981). Evaluating experienced teachers in its current form, as reflected in the sample that I have examined, seems merely a bureaucratic control of teaching. The watchful eye of the inspector is now the bureaucratic formality of the administrator.
References


Callahan, R. *Education and the Cult of Efficiency: A Study of the Social Forces that have Shaped the Administration of the Public Schools.* Chicago: University of Chicago, 1962.


Houston, S. E. “Politics, Schools and Social Change in Upper Canada”. The Canadian Historical Review S3, (3), September 1972.


Wilson, J. “The Pre-Ryerson Years.” In N. McDonald and A. Chaiton (Ed.), *Egerton Ryerson and His Times.* Toronto: Gage, 1977.


Appendices
# The Board of Education for the City of

## TEACHER REPORT COVER SHEET

(to be attached to all copies of Teacher Report)

<table>
<thead>
<tr>
<th>Teacher:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>School:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject:</td>
<td>Grade:</td>
<td></td>
</tr>
<tr>
<td>Date(s) of Visits(s):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date(s) of Pre-Conference(s):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date(s) of Post-Conference(s):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Assessment:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher's Comments:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have read this report and have received a copy:</td>
<td>Date:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teacher's Signature</td>
<td></td>
</tr>
<tr>
<td>Evaluator's Signature</td>
<td>Position</td>
<td>Date:</td>
</tr>
<tr>
<td>Principal's Signature</td>
<td>Date:</td>
<td></td>
</tr>
<tr>
<td>Distribution after teacher/evaluator/principal signatures:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associate Superintendent - White</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher - Yellow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principal - Pink</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P&P/4.04/cc
## COMPONENTS OF THE TEACHER REPORT

- Information & Signatures Specified on the Report
- Observation
- Areas for Growth
- General Assessment

## CRITERIA: STANDARDS OR PRINCIPLES FOR MAKING JUDGEMENTS

The evaluation should include the following criteria as appropriate to the individual and the individual's teaching area:

### CLASSROOM TEACHING PERFORMANCE

- **Planning:**
  - planning and preparation
  - organization

- **Instructional Methodology:**
  - knowledge of current teaching/learning theories
  - use of clear objectives and expectations
  - teaching techniques
  - attention to different learning styles
  - questioning techniques
  - correction or remediation

- **Management of Learning Environment:**
  - organization
  - class management
  - use of routine
  - communication skills
  - student involvement
  - teacher/student relations

- **Student Evaluation:**
  - alignment with teaching activities
  - diagnostic techniques
  - formative techniques
  - summative techniques
  - reporting to parents

- **Subject/Program Expertise:**
  - knowledge of subject/program
  - knowledge and application of underlying principles of the subject/program

The evaluation may include the following criteria:

### CONTRIBUTION TO TOTAL SCHOOL LIFE

- extra curricular involvement - student, staff
- leadership within the school
- curriculum development
- supports school policies
- supports school objectives
- assumed duties outside the classroom
- peer inter-relationships - colleagues, administrators, support staff

### PROFESSIONAL ACTIVITIES & PROFESSIONAL GROWTH

- knowledge of system philosophy, expectations, and program
- evidence of growth
- keeps abreast of new development - participation in professional association

### ADMINISTRATIVE DUTIES & LEADERSHIP

- role in school/department policy
- organizational skills

## GENERAL ASSESSMENT*

*GENERAL ASSESSMENT RATINGS MUST BE SHOWN FOR ALL PROBATIONARY TEACHERS AND FOR PERMANENT TEACHERS WHOSE WORK IS UNSATISFACTORY. OTHERS ARE OPTIONAL.

### FOR PROBATIONARY TEACHERS ONLY:

- **SATISFACTORY:** Meets Board standards.
- **UNSATISFACTORY:** Fails to meet Board standards.

**NOTE:** A probationary teacher requires a general assessment of "satisfactory" in order to qualify for a permanent contract.

### FOR PERMANENT TEACHERS ONLY:

- **EXCELLENT:** Exceeds Board standards on a regular basis.
- **VERY GOOD:** Meets Board standards on a regular basis.
- **SATISFACTORY:** Meets most of the Board standards on a regular basis.
- **UNSATISFACTORY:** Fails to meet Board standards.
The Board of Education for the City of
TEACHER REPORT

<table>
<thead>
<tr>
<th>Teacher:</th>
<th>Report Page Number</th>
<th>___ of ___</th>
</tr>
</thead>
<tbody>
<tr>
<td>School:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Date: ____________________  Principal's Initials: ____________  Teacher's Initials: ____________

Distribution:  (Associate Superintendent - White)  (Teacher - Yellow)  (Principal - Pink)

P&P/4.04/cc