



uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTDOCTORALES



FACULTY OF GRADUATE AND
POSTDOCTORAL STUDIES

Aletha Adell Hay

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Education)

GRADE / DEGREE

Faculty of Education

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

The Mediating Influence of Internal Accountability Systems on Teachers' Assessment Practices

TITRE DE LA THÈSE / TITLE OF THESIS

Marielle Simon

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Martin Barlosky

Brad Cousins

Lorna Earl

Colla Jean MacDonald

Gary W. Slater

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

The Mediating Influence
of
Internal Accountability Systems
on
Teachers' Assessment Practices

Aletha Adell Hay

Thesis submitted to the
Faculty of Graduate Studies
University of Ottawa
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy in Education

Faculty of Education
September 2005

© Aletha Adell Hay, Ottawa, Canada, 2005



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-494-10974-2

Our file Notre référence

ISBN: 0-494-10974-2

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

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

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Acknowledgements

In completing this dissertation, I must acknowledge the many people who have encouraged and supported me throughout this process from its beginning to its conclusion.

I would first like to extend my thanks and admiration to my thesis advisor, Dr. Marielle Simon, a committed professional and invaluable mentor. I would also like to thank the members of my thesis committee, Dr. Bernie Andrews, Dr. Bradley Cousins, and Dr. Colla Jean MacDonald. Their perspectives and suggestions provided helpful direction in the development of this thesis. I admire your abilities and enjoyed working with you. I especially appreciate the contribution of Dr. Martin Barlosky who joined my committee in the late stages of the writing but whose advice was most pertinent.

To my family, I owe heart-felt appreciation for the unconditional love and unfailing support. To my husband Ron, I value your respect for my educational pursuits. To my sons Graeme, Andrew and Alan, I am grateful for your patience and unswerving confidence in my efforts to accomplish this project.

Abstract

This collective case study explores the nature of internal accountability systems in two elementary schools situated within a provincial-level accountability agenda using large-scale performance testing to ensure both school accountability and professional learning. The purpose of this mixed model study was threefold: first, to explore the components and strength of internal accountability systems and the interactive dynamics with the external system; second, to identify key environmental factors that impacted the school-based systems; and third, to reveal the influence these internal forums had on teachers' assessment practices. Multiple theoretical lenses informed the conceptual framework used to analyze the policy to practice connections in the restructuring educational system.

The study findings reveal several observations. First, internal accountability systems should possess five components including aims, information on school performance, a standard of achievement, person(s) who judge school performance, and mechanisms to guide and monitor teachers' practices. The presence of all five components suggests that strong internal accountability systems were in place in the two schools. Second, two environmental factors impacting these systems were large-scale performance tests and professional leadership. There is indication that the design of the testing program minimized unintended negative effects linked with public reports and that school leadership was instrumental in mobilizing support for school improvement and professional learning. Third, the study also describes the relationship between the internal and external systems nested within the province's standards-based reform initiative. The findings unpack the compatibility between the instruments and strategies

of the school-level systems and those of the state-level system, lending support for the proposition that external oversight contributes to school-based ownership for proactive change. Finally, the study findings illuminate how internal accountability systems mediated teachers' practices and supported the development of collegial assessment practices. The findings lend support for the position that the dual aims of accountability and professional learning are enabled when internal accountability systems are facilitated by a professionally oriented external policy environment.

Table of Contents

Acknowledgements.....	ii
Abstract	iii
Table of Contents.....	v
List of Figures.....	viii
Introduction	
Statement of the Problem.....	1
Overview of the Thesis.....	7
Chapter 1	
Review of the Literature	
Introduction and Overview.....	9
Accountability Framework.....	10
Accountability defined.....	11
Dimensions of an accountability framework.....	11
Accountability in context.....	27
Internal accountability systems.....	29
Assessment in Education.....	35
Paradigm shift in assessment.....	35
Assessment defined.....	37
Classroom assessment	39
Assessment community.....	46
Theories of System Change.....	50
Background issues.....	51
The change process in Ontario.....	54
Research Questions.....	62
Conceptual Framework.....	64
Framework for analyzing systemic change.....	64
The conceptual model.....	67
Overview of the Chapter.....	78
Chapter 2	
Methodology for the Study	79
Participants	81
Province	82
District	85
School	87
Teachers	91
Data Collection	92
Instruments	92
Procedures	95
Data Analysis.....	98
Trustworthiness	102
Chapter 3	
The Case of Southview School.....	106
Internal Accountability System at Southview School.....	106

	Strength of Southview School's Internal Accountability System.....	107
	Aims	107
	Information on school performance.....	108
	Standard of achievement.....	111
	Persons who judge school performance.....	113
	Mechanisms to guide/ monitor school practice.....	116
	Factors Impacting Internal Accountability at Southview School.....	129
	EQAO testing.....	130
	Leadership.....	134
	Political constraints.....	139
	The Relationship with the External Accountability System.....	140
	Compatibility of aims of accountability.....	140
	Compatibility of information on school performance.....	142
	Compatibility of standard of achievement.....	143
	Compatibility of persons to judge school performance.....	143
	Compatibility of accountability instruments and strategies....	143
	The Nature of Southview School's Internal Accountability.....	147
	Influence of Internal Accountability on Assessment Practices	151
	Classroom Assessment	151
	Purpose.....	151
	Criteria for selection.....	157
	Assessment methods.....	159
	Feedback.....	161
	Dynamics of School Assessment Practice.....	163
	Assessment goals at Southview School.....	163
	Assessment instruments at Southview School.....	165
	Assessment methods at Southview School.....	165
	Personnel involvement in assessment.....	166
	Values	168
Chapter 4	The Case of Northside School.....	171
	Internal Accountability System at Northside School.....	172
	Strength of the Internal Accountability System.....	172
	Aims.....	172
	Information in school performance.....	174
	Standard of achievement.....	176
	Persons who judge school performance.....	179
	Mechanisms to guide/ monitor school performance.....	181
	Factors Impacting Internal Accountability at Northside School.....	193
	EQAO Performance tests.....	193
	Leadership.....	196
	The Relationship with the External Accountability System.....	198
	Compatibility of aims of accountability.....	198
	Compatibility of information on school performance.....	199
	Compatibility of standard of achievement.....	200
	Compatibility of persons to judge school performance.....	200
	Compatibility of accountability instruments and strategies....	200

	Nature of Northside School's Internal Accountability.....	204
	Influence of Internal Accountability System on Assessment Practices...	209
	Classroom Assessment	209
	Purposes.....	209
	Criteria for selection.....	214
	Assessment methods.....	216
	Feedback.....	218
	Dynamics of School Assessment Practices.....	220
	Assessment goals at Northside School	221
	Assessment instruments at Northside School	221
	Assessment methods at Northside School.....	222
	Personnel involvement in assessment.....	224
	Values	225
Chapter 5	Interpretation and Discussion.....	229
	Cross-Case Findings.....	229
	The Strength of Internal Accountability Systems.....	230
	Factors.....	242
	Relationship with External Accountability.....	255
	The Nature of Internal Accountability Systems.....	270
	Influence of Internal Accountability Systems	275
	Discussion.....	290
Chapter 6	Conclusion.....	305
	Contributions and Highlights.....	306
	Limitations.....	316
	Implications for Research.....	319
	Implications for Practice.....	323
References		329
Appendixes		
	Appendix A- Informed Consent (teacher).....	341
	Appendix B- Informed Consent (school leader).....	344
	Appendix C- Informed Consent (principal).....	347
	Appendix D- Informed Consent (district officer).....	350
	Appendix E- Informed Consent (EQAO officer).....	353
	Appendix F- Interview Protocol 1.....	356
	Appendix G- Transcription /Data Retrieval codes.....	360
	Appendix H- Figures (display of full proofs).....	363

List of Figures

Figure 1	Integrated Model of Internal Accountability Systems.....	70
Figure 2	Factors Impacting Internal Accountability Systems	74
Figure 3	Dynamics of School Assessment Practice	77
Figure 4	Cross-case Comparison of Internal Accountability Systems.....	232
Figure 5	Cross-case Analysis: Factors Impacting Internal Accountability.....	243
Figure 6	Cross-case Analysis: Nested Context of Accountability Policy.....	256
Figure 6 (Continued)	Cross-case Analysis: Compatibility of Accountability Mechanisms.....	257
Figure 7	Cross-case Analysis: School Response to Accountability.....	272
Figure 8	Cross-case Analysis: School Assessment Practices	277
Figure H1	Dimensions of an Accountability Framework.....	363
Figure H2	Avenues of Influence: Factors Influencing Internal Accountability Systems	364
Figure H3	Classroom Assessment	365
Figure H4	Characteristics of School Assessment Practices	366
Figure H5	Continuum of Approaches to Accountability.....	367
Figure H6	Southview School Internal Accountability Components	368
Figure H7	Factors Impacting Southview School's Internal Accountability System	370
Figure H8	Relationship Between External and Internal Systems	371
Figure H9	Southview School Response to Accountability Aims	373
Figure H10	Classroom Assessment	374
Figure H11	Dynamics of Southview School Assessment Practices.....	375
Figure H12	Northside School Internal Accountability Components.....	376
Figure H13	Factors Impacting Internal Accountability at Northside School.....	378
Figure H14	Relationship Between External and Internal Systems.....	379
Figure H15	Northside School Response to Accountability Aims	381
Figure H16	Classroom Assessment	382
Figure H17	Dynamics of Northside School Assessment Practices.....	383

INTRODUCTION

Studies of systemic reform initiatives in education over the past decade report limited success in terms of meaningful change in teachers' practices and school performance. Issues related to policy implementation, change processes and professional learning are all identified as factors hindering school improvement. Demands for accountability led to closer examination of accountability processes, many of which involved some form of testing. It seems that the anticipated impact of external accountability systems is hampered by unintended negative effects associated with large-scale testing and by less than successful implementation strategies to enhance school capacity. These drawbacks have invited exploration of a new conception of accountability that takes into consideration the dynamics of school organizations and the perspectives of the educators within them.

Emerging from studies of accountability systems is a conceptual model, one that is school-based and connected with high instructional capacity and effective school programming. Proponents of this model suggest a supportive rather than a controlling role for external systems, one that aims to enable strong school-based responsibility by providing objective measures of school performance and by encouraging change in teachers' practices through professional learning. The underlying premise appears to be a joint ownership for accountability that balances external oversight with internal responsibility for school improvement. This premise implies that professional ownership is integral to effective school improvement.

Traditionally, accountability systems were external, compliance-oriented processes imposed on schools. These systems first measured school success in terms of compliance with policy and later with measures of student achievement. Currently, accountability is explained as processes by which districts and states ensure that schools meet their goals, but this view is further elaborated by others to mean those commitments, policies and practices that foster the goal of student exposure to good instruction and informed educational practices (Leithwood, Edge & Jantzi, 1999; Leithwood & Earl, 2000; Darling-Hammond & Ascher, 1991). These definitions suggest differing perceptions of what schools are accountable for, with one view outlining that schools are accountable for student achievement and, in contrast, a second view articulating that schools are accountable for quality instructional programs. The logic of the second view implies that since many non-school factors contribute to student achievement; schools are responsible for only those factors within their control, specifically, the quality of the instructional program. These divergent perspectives likely explain the debate and uncertainty surrounding accountability policies in current system-wide change initiatives and seem to reflect the same paradigm shift inherent in the standards movement as schools move from culture of testing to a culture of assessment.

The impact of state-level accountability programs to sustain change in teachers' instructional practices through external testing is described and disputed in the literature (Chrispeels, 1997; Darling-Hammond, 1990; Firestone, Mayrowitz & Fairman, 1998;) and found wanting for several reasons. These include lack of coherence between: the tests and the curriculum expectations, the resources for testing and for professional learning, and the performance tests and teachers' practices.

Newman, King & Rigdon (1997) studied 20 restructuring schools nested in educational systems possessing strong, inconsistent, and negligible external accountability. From the study findings, these researchers learned that strong accountability can exist *internal* to a school and be independent of district and state systems. In comparing strong external systems versus strong internal systems, they reported that internal accountability was more likely than with strong external accountability to be associated with high organizational capacity. In their conclusion, these researchers argue that three issues kept the promise of the popular theory of strong external accountability from working. These issues are summarized as controversy related to standards, incentives and governance issues; insufficient effort to support school capacity with appropriate human, technical and social resources; and the failure to recognize the role of internal accountability at the school level.

To examine accountability systems, Newman and his colleagues (1997) identified four central features of an accountability system as being: information on schools performance, standards for measuring achievement, person(s) who judge the level of success, and consequences for success or failure. They proposed that strong accountability systems possess all four components and contrasted weak systems as those with standardized testing but no standards and no consequences. Mid-range systems possessed several but not all of the components. In this model, the source of the consequences was deemed irrelevant and could be externally required, internally generated or a combination of both.

Another empirical work examined internal accountability from the perspective of the individual school 'nested' within differing external policy environments. In their

multiple case study of how schools think about accountability, Abelman & Elmore (1999) explored school-level conceptions of teachers' responsibility and accountability processes embedded in the day-to-day operations of the schools. These researchers proposed that schools developed normative internal structures (routines and professional perspectives) that were relatively immune to external influence. They suggest that school response to external accountability is a function of the degree of alignment between the school's internal accountability mechanism and the requirements of the external system, as opposed to the details of the external system. As the schools selected for the study precluded any reference to particular external accountability approaches, the nature of the relationship between the school structures and the external requirements remain problematic. However, this work elucidates a somewhat different conception of internal accountability, one that is framed by socio-cultural relationships such as conceptions of responsibility, shared expectations, and accountability mechanisms.

These empirical works and other reports (Cohen, 1995; Darling-Hammond, 1991; Earl & LeMahieu, 1997; Knapp, 1997) caution about the limitations of systemic change based on external policy-driven accountability systems and seem to highlight the reasons why policymakers should focus on encouraging the type of internal accountability associated with professional learning and organizational capacity. The arising question remains, what role should external policy play in system-wide accountability initiatives so integral within contexts of systemic change. If accountability is demanded by the public, how can the benefits of strong internal accountability be facilitated by an active external system. Although they do not

articulate support for uniform testing, Newman et al. (1997) propose that there is a substantive role for external agencies and that external contribution to school-based accountability can take the form of providing concrete examples of high performance standards, models of different approaches to performance assessments, and reliable ways of evaluating student performance. The next logical step would seem to be a consideration of what sponsorship and support from external agencies can facilitate such an effective, collective enterprise, especially in contexts where the public demands accountability.

New accountability systems are emerging that recognize schools are responsible for improving the quality of instructional programs (Fuhrman, 1999) and differ from traditional approaches in many ways. All seemed to be based on some form of testing with the emphasis on experimenting with performance assessments. Many appear to be reliant on tangible consequences in the form of reward or sanctions; and there seems to be controversy around the question of whether these new accountability systems should focus on “getting teachers’ attention” or on “supporting school capacity-building”. As Fuhrman concludes in her CPRE report, these new systems appear insufficient to improving student achievement in the absence of support for capacity building.

The possibility of a state-level accountability framework to design and implement large-scale testing that addresses Fuhrman’s two concerns, getting teachers’ attention and capacity building, is being tackled in the province of Ontario by the EQAO (Earl & Torrance, 2000). Based on theories of change (Hargreaves, 1994) and new theoretical directions in assessment and evaluation (Biggs, 1995; Earl & LeMahieu, 1997; Gipps, 1994; Linn, 2000), the EQAO has implemented a large-scale testing

program designed to address the dual aims of *assessment as accountability* and *assessment as professional learning*. Its requirements and recommendations reflect knowledge of change processes and professional learning theory and, as reported by Earl & Torrance (2000), seeks to address the limitations of earlier models of accountability reported in the literature. The testing program reflects the perspective that internal and external accountability systems may be discrete but complementary under the proper conditions that facilitate *assessment conversations* (Earl & LeMahieu, 1997) between school and community stakeholders.

To explore whether the intentions and policy requirements of a standards-based accountability program support school-based accountability and enhance professional learning warranted attention. The challenge of describing and analyzing policy to practice connections in contexts of systemic change is complex because, as Knapp (1997) explains in his review of the literature, no single theoretical perspective can address all outstanding issues. Such connections are best understood using an analytical framework informed by theories from policy implementation, innovation and change, and professional and organization learning. He also reminds readers that systemic reform thinking anticipates that aligned guidance and support influences teachers. Then he identifies four essential elements necessary to capture such connections as the target of the policy, the policy instrument, the avenues of influence, and the context and conditions that mediate policy.

Little is known about the nature of internal accountability systems in educational contexts where active external accountability programs aim to ensure accountability by getting teachers' attention and aim to promote professional learning by using

performance tests that model curriculum-linked tasks and assessments. The findings of this collective case study present tentative suppositions related to the nature of internal accountability and the influence of such school-based systems to mediate teachers' practices. The study findings have implications for further research on internal accountability systems and have implications for enhancing teachers' practices.

Overview of the Thesis

This report on the collective case study is organized in six chapters. Chapter 1 reviews the literature used to frame the theoretical propositions and guiding questions. This chapter also describes Ontario's innovative accountability program and articulates the aims of the Education Quality and Accountability Office [EQAO]. Next, the conceptual framework is outlined and the conceptual model is operationalized using concepts from the literature. Following this, the intent of this research is further elaborated. Chapter 2 explains the qualitative methods and mixed model design used in the study. The selection of the case study participants is outlined and the external policy context is framed. Data collection procedures are then reported and the data analysis detailed. The chapter ends with a discussion of the trustworthiness of the study design. Chapters 3 and 4 present the case study findings for Southview School and Northside School respectively. Each chapter is organized by the format of the research questions, first describing the nature of the internal accountability systems and then exploring the influence of these school-based systems on teachers' assessment practices. Chapter 5 presents the cross-case analysis and interpretation, followed by a discussion of the findings. Chapter 6 reports on the highlights of the study and its contribution to theory.

This is followed by a statement of the limitations. The chapter concludes by presenting the implications for research and for practice.

CHAPTER 1

LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

This chapter provides background to the theory and the issues that frame the purpose and the intent of this study of internal accountability systems. The review outlines the dimensions of accountability and the elements that comprise each dimension. The review also provides a rationale for each of the research questions, summarizing what the literature reports and then clarifying the assumptions that limit the focus of this qualitative inquiry on policy-to-practice connections. The conceptual framework for the study is described and its elements are operationally defined in order to identify the components of the integrated model of an internal accountability system. An overview of this chapter and a preview of the next chapters conclude this literature review and conceptual framework.

Overview of the Research and Theoretical Framework

This study of internal accountability systems, their relationship with the external system and their influence on teachers' assessment practices relies on research in three areas. The first is research in the foundational work associated with the conception of accountability and the theoretical perspectives that inform and define the dimensions of existing approaches found within education systems internationally. This area of research helps to frame the six dimensions of an accountability framework: mode, aims, form, factors and organizational context, and conditions. This review is useful in recognizing different approaches and in analysing the provincial context in which this study is situated.

The second area of research reviewed deals with the field of measurement and evaluation, specifically, the debate and promise of performance assessment and its role within large-scale testing programs. The role and the validity of such assessments to both guide educational decision-making and inform educational practice is examined.

The third area of research is the field of system change in education. This part of the literature review differentiates reform approaches from restructuring initiatives. It provides background related to the system change movement in the province of Ontario and details the aims and structure of its innovative framework. This research domain provides guidance on the policy to practice connections that are examined in the implementation of the large-scale assessment program associated with the external accountability system. It provides an outline of the comprehensive theoretical (analytical) model used to develop the conceptual model for this collective case study on internal accountability systems in Ontario schools. This field justifies the focus of the study on school structure and professional interactions, specifically the nature and influence of internal accountability systems on teachers' assessment practices.

From these research areas, propositions are formulated and questions for the study are theoretically and methodologically framed. The following review of the literature contributes to the study's conceptual framework and provides the basis for both the operational definitions of internal accountability systems and the research questions.

Accountability Framework

This section of the review begins by presenting a definition of accountability as it relates to educational organizations. Next, the dimensions of an accountability

framework are explored and after that, what the literature has to say about the context and the internal accountability systems of schools.

Accountability Defined

Accountability is a term with many meanings. Two specific definitions for educational accountability found in current literature both clarify that the term means actions related to achieving goals but differ somewhat in what the nature of the goals should be. From a professional (organizational) learning perspective (Darling-Hammond & Ascher, 1991), accountability is described as commitments, policies, and practices that foster student exposure to good instructional practices, reduce the likelihood of harmful practices, and involve self-corrective procedures on the part of the provider of the service. In contrast, the definition developed by Rothman (1995) and foundational to Leithwood, Edge and Jantzi's categorization of accountability approaches is "the process[es] by which school districts and states attempt to ensure that schools and school systems meet their goals." (1999, p.189) Understood in current standards-based reform agendas (Leithwood & Earl, 2000), this goal is student achievement because it is the criterion upon which to judge the value of an educational system. For the purposes of this study, both goals are valid because the two fundamental aims of education to provide a quality instructional program and to enhance student achievement are integral parts of the teaching-learning dynamic.

Dimensions of an Accountability Framework

The accountability framework adopted for this study has seven dimensions including the conception (underlying ideas and assumptions), the approach or mode, the

aims and consequences, the instruments and strategies comprising the form, the factors that make up the conditions that influence success, the locus of authority on control of schools (system level), and the depth and breadth of organizational involvement (see Figure H1). These dimensions constitute parts of an accountability system that, in theory, possesses specific cohesively elements linked to underlying beliefs and assumptions about the roles and responsibilities of schools. Each dimension is addressed as a sub-section below and aims to explain the elements based on current literature.

Conception of Accountability

In the work *Educational Accountability: State of the Art* (Leithwood, Edge and Jantzi, 1999), the authors refer to earlier works by R. Wagner as they discuss five features of that comprise a conception of accountability. These features are best explained as responses to the following key questions:

1. What level of accountability is to be provided?
2. Who is expected to provide the account?
3. To whom is the account owed?
4. What is to be accounted for?
5. What are the consequences of providing an account? (p.13)

The underlying beliefs and ideas about what goals, processes and sources of authority are legitimate within an educational system help to understand the conception of accountability that is being used to develop an accountability system. Together responses to these questions first identify the assumptions related to accountability and second, structure the development of the approach to be employed.

The first question--*What level of accountability is to be provided?*-- speaks to the level or intensity of the accounting that must be made. The level of accountability may

vary from a simple report or detailed explanation, to a full justification for decisions, actions and outcomes.

The second question--*Who is expected to provide an account?*—seeks to explain what person or group are held responsible and reflects in part what is considered as a legitimate goal. As Leithwood and Earl (2000) point out, it is legitimate to hold a teacher or principal accountable for the most productive use of resources to reach school goals but not legitimate for them to be held solely responsible for student achievement because many of the factors that contribute to achievement are outside the influence of school personnel.

The third question--*To whom is the account owed?*-- addresses the issue of entitlement. If the school is responsible for providing a quality instructional program, the accounting then relates to professionalism. It is then incumbent that both school administrator and teachers provide an account of what they do to the public.

The fourth question--*What is to be accounted for?*-- focuses on the fundamental core of education, the welfare of the student. This welfare refers to academic achievement but may incorporate additional elements such as organizational effectiveness, professional knowledge and skill, best practices, moral behaviour, or skilful use of organization change processes (program evaluation, planning).

The fifth question--*What are the consequences of providing an account?*-- delineates what consequences, if any, are associated with making an account. The nature of the consequences is directly linked to the aims of the accountability approach and this is, in turn, linked to the underlying beliefs and assumptions.

These five questions are pivotal to understanding the ideas held by policymakers and stakeholders about matters of accountability. As explained by Leithwood et al.

(1999), responses to the five dimensions of accountability vary depending on the beliefs that policymakers and the public hold about the responsibilities of schools. In the literature (Corcoran & Goertz, 1995; Fuhrman, 1999; Fullan, 1998; Harris, 2000; Hatch, 1998; Knapp, 1997; Leithwood & Earl, 2000), conceptions of accountability range from traditional rationale management models to models incorporating empowerment of decentralized change agents. The responses also vary depending upon stakeholder perspectives about how much change is needed in schools and what processes are best able to bring about that change.

Approaches to Accountability

The second dimension of accountability involves its approach. In the classification scheme reported by Leithwood et al. (1999), four approaches to accountability are listed – market competition, decentralization, professionalism, and management approaches. These approaches reflect different combinations of the five features discussed above, but they also represent differentiation based on the instruments and strategies used as change mechanisms. This schema for alternate approaches is further developed by Leithwood and Earl (2000) because of the emerging emphasis on moral sources of authority and control in restructuring systems. These authors justify professional approaches rooted in principled, ethical professional practice as the source of authority and legitimacy most likely to explain the effectiveness of internal accountability mechanisms and the uniqueness of school response to external accountability.

Approaches to accountability that allocate authority to centralized decision-makers have been found wanting in terms of sustainable change in teaching practice as in depth case studies of policy implementation have shown (Chrispeels, 1997; Cizek,

Fitzgerald & Rachor, 1996; Firestone et al., 1998). Other theoretical works (Fuhrman & O'Day, 1996) propose that an appropriate selection and balance of rewards is the solution to ensuring educational effectiveness. However, much of the investigation of teacher practice (Louis & Marks, 1998) affirms that it is rooted in beliefs about responsibility to students and understanding of how students learn. Instructional change that is sustainable must address teacher learning and must involve the type of professional development that is situated and collegial, if new practices are to be rooted in what teachers think and do in their classrooms. This second conception of accountability contrasts with the former because professional values and commitment are less linked to extrinsic rewards but tied to the intrinsic satisfaction of being an effective and student-oriented teacher. These latter approaches contrast, in part, with the Adams and Kirst's (1996) model because the schema developed by Leithwood et al. (1999) de-emphasizes legal and bureaucratic sources of accountability and control. One must ask if the emerging professional-based models of accountability will overcome the limitations associated with high stakes systems (Firestone et al., 1998; Mintrop, MacLellan & Quintero, 2001; Newman et al., 1997) when the focus is on effective teaching and good instructional practices.

Accountability frameworks use several devices to guide and control system actions. According to McEwen's (1995) review, a good accountability framework will use a variety of strategies and these may include system changing (restructuring), capacity building (teacher learning), inducements (rewards, sanctions), and mandates (requirements). However, accountability is a term that has increasingly become synonymous with testing (Linn, 1998) and the establishment of large-scale testing programs is a powerful policy lever in education.

Aims and Consequences of Accountability

The third dimension focuses on the aims and consequences of accountability. According to Fuhrman (1999) in her position paper on the evolving definition of accountability the conception of accountability in the United States has been influenced by the same paradigm shift that created the standards movement. She explains emerging accountability approaches and their systems as a shift from traditional (rational-structure) thinking to a new perspective that possesses seven features: a) a focus on school performance; b) schools are the unit of improvement; c) continuous improvement strategies; d) school inspections and reviews; e) school certification and rating scales; f) public reporting; and g) consequences attached to performance levels. Her interim report on this *New Accountability* describes a shift from a passive compliance strategy to an accountability strategy that involves public reporting and dependency on issues of motivation and capacity. The emerging aims in this conception of accountability does not appear to own capacity issues from a professional learning perspective, specifically, aspects such as teacher professional knowledge, principal leadership, and professional community. This is short-sighted in view of the current qualitative and quantitative research findings (Cizek et al., 1996; Corcoran & Goertz, 1995; Daniel & King, 1998; Grant, Peterson & Shojgreen-Downer, 1996; Hall & Harding, 2002) that report the need for quality pre-service preparation and in-service professional development, while highlighting the value of professional networks and collegial inquiry into best practices and school improvement.

In a theoretical essay based on observations of three empirical studies of ambitious educational reforms, Van der Vegt, Smyth, and Vanderberghe (2001) contend that school level response to policy changes created organizational issues that triggered

teacher concerns. One example of this related to the addition of professional development to the organization of staff meetings, raising teacher concerns about the perception of their competency and effort. These researchers suggest that interplay between issues and concerns can be addressed by using school procedures such as discussing school goals, sharing best practices, and implementing collegial planning. These authors propose that policy implementation use processes to resolve teacher concerns related to inclusion, competence, influence and fairness. Their position implies such processes that resolve teacher concerns because they involve teachers as professionals. This approach aligns with Fullan (1998)'s position that governments must:

Counterbalance or integrate these (compliance instincts) with capacity building policies, incentives and supports - things that governments are not traditionally inclined to do...appreciating that policies on student assessment can and should serve both the accountability function of making everyone aware of how well students are doing, and the implementation function of developing strategies to make improvements based on the results. (p.6)

These works appear to add the element of problem-coping through collegial dialogue to Fuhrman's (1999) seven features of the new accountability movement. It would seem that professional concerns must be considered part of the change equation and that strategies that build organizational capacity and incorporate a professional learning component within the accountability process are worth consideration. Accountability, the literature suggest, should aim to both require an account of school performance and provide opportunities to enhance professional learning.

Re-conceptualizing this new accountability as having two complementary but different purposes is proposed. Whether or not rapport between the intentions of administrative policy and the school level response can be coordinated has implications for research. What strategies are needed that will balance the interests of both aims, address issues of too much or too little structure, and satisfy top-down desires with bottom-up inclinations? Fullan (1998) proposes the deliberate formulation of combined strategies, trying out and learning from these trials in order to refine and strengthen the system capacity to mobilize school and state forces in concert (p.7). How these dual purposes play out in real schools and classroom warrants examination. The question that arises is what combinations of instruments and strategies are compatible with these dual purposes and will they complement the responsibilities of policymakers on the one hand and respect the professionalism of teachers on the other. A study that examines the strength of existing internal accountability in a system of external accountability that recognizes its importance is warranted.

Form of Accountability

The fourth dimension has to do with accountability system *instruments* and *strategies*. These fall under the form of accountability and have been broadly described and categorized in the literature (McEwen, 1995; Lapointe, 2000). For the purpose of this study, *instruments* associated with systemic change are the requirements that guide school practice. These may include: a) curriculum frameworks, b) achievement standards, c) large scale testing programs, d) standardized report cards, e) school restructuring and governance, and f) professional certification. Of these, large scale testing programs is the prominent instrument for measuring and reporting on system performance and is therefore the key element of most frameworks.

Associated with such instruments are specific *tools* that guide but also monitor how the process of implementation is occurring by providing feedback. Such tools may include: a) strategic and improvement planning, b) indicator systems, c) testing, d) evaluations, e) opinion surveys, and f) public reports. Inherent within these mechanisms is feedback that, dependent on the approach to accountability, is akin to consequences.

Large-scale testing programs are a popular form of accountability for many reasons. Linn (2000) explained that they are a comparatively inexpensive option for maximizing system efficiency, are externally mandated, can be implemented quickly, and have visible results. The emergence of *performance-based assessments* (Linn, 1993; Herman, 1997) as vehicles for accountability systems indicates state level interest in establishing standards and getting teachers to focus on student achievement.

In a review of testing programs, Herman (1997) notes that “what is tested is what is taught” (p.203). She found that teachers: a) showed increased attention to student achievement, b) became more familiar with models of good instructional activities, c) used student scores as feedback on instructional program and teaching effectiveness, and d) encouraged parent involvement in student learning when large scale tests were used to track student achievement. It follows that if teachers do pay attention to the content and processes of such tests, a well-designed test that reflects the type of instruction and assessment desirable in classrooms can provide a model for teachers about what is wanted in terms of curriculum, assessment and pedagogy (Resnick, L. & Resnick, D.,1992). Firestone et al.(1998) cautioned that high-stakes associated with large-scale performance tests may motivate change in teacher practice but must be linked with opportunities to learn about performance assessments in different subject areas.

However, with such popular reasons for testing, Linn (1993) cautions that the nature and design of performance tests do vary and issues of generalizability and validity with performance tests do need to be addresses. According to Haertel (1999), who resists the consequential validity argument in its entirety, psychometricians should ask questions about educators' perceptions on the usefulness of test scores as adequate measures of what they are accomplishing and he also goes so far as to suggest that teachers may help identify weaknesses in the tests themselves. Other validity issues related to the unintended negative effects of testing (Cibulka, 1991; Haertal, 1999, Shepard, 1997) are the comparison of schools based solely on a single test result, the public right to know versus privacy issues, teacher retrenchment and the diversion of public attention to testing away from the root of the problem – quality instruction.

Large-scale tests are one vehicle for addressing public demand for accountability (Earl & LeMahieu, 1997; Mawhinney, 1998) and tremendous resources are expended in system development of large-scale test programs involving performance assessments. Such testing programs can signal what is important for students to learn and can model effective assessment strategies (Herman, 1997). As such, these types of test programs are able to make important contributions in terms of feedback and motivation, but tests alone cannot break the need for change down into manageable stages or steps that guide change in teachers' practices at the school level.

An examination of performance-based tests to inform teacher practice and to promote system transparency is warranted (Cibulka, 1991; Earl, 1995), but more importantly, a closer examination of how test-based accountability informs teacher practice is needed to determine if it triggers school-based mechanisms that influence teachers' instructional and assessment practices. This possibility is considered next in

light of what the literature says about factors that influence success of reforms and then what is known about those dimensions of accountability that frame the context and conditions for change.

Accountability *mechanisms* (Abelmann & Elmore, 1999) are defined as those formal and informal ways that a person(s) makes an account of their actions. School accountability mechanisms are many and varied, but they may be linked to specific kinds of feedback or consequences that come about as a result of the account taking, in other words meeting or not meeting established expectations. Feedback associated with accountability mechanisms can trigger change at the professional and organizational level.

Factors Influencing Accountability Frameworks

Factors influencing accountability frameworks represent the fifth dimension of accountability systems. Research (Claussen, 2001; Knapp, 1997) suggests several factors influence policy to practice connections in schools. Knapp reports that multiple avenues of influence exist, but identifies five elements common in policy to practice connections: a) *policy requirements*, b) mobilization of *supportive actors*, c) *Professional Networks* as the environment of professional ideas, d) *resources and supports*, and e) *professional qualifications* based on certification requirements and Standards of Practice (see Figure H2). Of these, Knapp argues that two represent missed opportunities: firstly, system investment in resources and support systems to sustain teacher learning, and secondly, changes in teacher education and certification. These avenues of influence may be enablers or hindrances depending upon the context and conditions, but all contribute to the *sphere of influence*.

The *policy requirements* of an accountability system may possess several elements depending upon the approach being employed. Four generic classes of alternate accountability instruments include mandates, inducements, capacity building, and system changing. McEwen (1995), in her report on accountability tools, expands on those requirements related to student learning and improved performance. For the purposes of examining requirements that impact accountability systems, the accountability mechanisms related to large-scale testing programs include: a) curriculum expectations and standards of achievement, b) large-scale performance testing and c) system reporting. In other words, the tools associated with accountability instruments discussed previously and directly involved in the act of making an account.

For the purposes of this study, the second avenue of influence is examined from the perspective of transformational leadership and moral authority. The *mobilization of supportive actors* such as superintendents, consultants and school administrators that press for change and encourage experimentation among school faculty is a conduit to inform school decision-making. Instructional quality can be strengthened when action is taken to organize professional interactions, align school professional development with school goals, promote socialization among the staff, and practice shared leadership (Young & King, 2002). These aspects of school leadership are closely examined in many current works. Studies of leadership have shown an evolving trend of co-joining school management and educational leadership. Educational leadership, originally based on work in the field of political studies and corporate management, must also appreciate the moral element inherent in public education.

Fullan's (2003) treatise on school leadership assumes moral authority is the pre-eminent form of authority governing school responsibilities. His work doesn't debate the

nature of the moral ecology but emphasizes the responsibility of schools to educate the young to be knowledgeable and self-actualized citizens in a democratic society. Transformational leadership, in his opinion, is not based on authority alone, but on a relationship based on mutual needs, aspirations and values.

The effectiveness of transformational leadership to facilitate teacher commitment to school change is reported in several empirical works (Barnett & McCormick, 2003; Fullan, 2003; Geijsel, Slegers, Leithwood & Jantzi, 2003; Young & Kings, 2002). In particular, the Barnett And McCormick's findings revealed that the relationship between the principal and the teachers was a pivotal factor in motivating teacher commitment to change. The trustworthiness and competence of the principal leaders was reflected in their concern for individuals evidenced by leadership behaviours which included "accessibility, encouragement, provision of structures and resource support and recognition" (p.70). Individual consideration was more than consideration and helpfulness; it was developmentally oriented because it balanced professional autonomy with the need for structure. These findings align with Young and King's (2002) multi-year, qualitative study of nine urban elementary schools where principal leadership focussed on professional development. Common themes in these studies include establishing trust and creating structures that promote teacher learning.

The mobilization of supportive actors at the school level is a significant factor in understanding policy-to-practice in restructuring schools. Other factors that impact teachers practices include as: a) professional networks, b) the availability of resources and supports, and c) the teacher qualifications related. Findings related to these remaining environmental factors are discussed briefly.

A third factor is the environment of professional ideas and involves *professional networks* and opportunities for professional development. This factor is directly linked to a fourth factor, *resources and supports*. Corcoran and Goertz (1995) highlight the debate about the type of capacity building needed in restructuring school systems. Their position defines the product of education as 'high quality instruction' and the means to achieve it the 'instructional capacity' of the school. Based on their review of the literature on organizational effectiveness and job performance, three sets of variables emerge as impacting organizational capacity: improving human capital, improving the quantity and quality of resources, supporting instructional collegiality, collaboration and cooperation. Spillane and Thompson (1997) reporting on a five-year study of instructional reform in nine school districts, identified these same three interrelated dimensions of local capacity. Findings of a sub-ordinate eighteen month project illuminated the significant impact of professional networks on increasing teacher disciplinary knowledge and changing practice.

These findings support earlier empirical work (Little, 1993; Peterson, McCarthy & Elmore, 1996) that reported changing teachers' practices was primarily a problem of learning, not of organization. Peterson and colleagues found school structures can provide opportunities for learning new teaching practices through professional dialogue and collaboration. Little's (1993) essay on teacher development in contexts of educational reform, first provides a synopsis of findings from five streams of reform and, secondly, identifies four effective alternatives to the training model. Her list included teacher networks, subject matter associations, school-university collaborations, and special institutes. The fifth factor, *Teacher qualifications* and expertise in the area of student assessment, is rated as a significant area in need of improvement in the literature

(Biggs, 1995; Black & Wiliam, 1998; Stiggins & Bridgeford, 1986) and this further supports that professional learning is important in contexts of educational change.

The factors presented here represent some of the conditions that comprise the *sphere of influence* that may supports change initiatives and build consensus and commitment to organizational goals (Knapp, 1997).

Locus of Authority

The sixth dimension of an accountability system is the *locus of authority*. The term *locus of authority on control of schools* (Knapp, 1997), as it is used in this study, refers to the authority to make decisions related to one's responsibilities. It perhaps can better be explained as responsible decision-making (Hargreaves, 1994). The term is based on the premise that it is the teachers, principals and other educators who must own the responsibility for being accountability for what they do, furthermore teachers must own the responsibility to evaluate their instructional as well as their students' performance. Inherent in this program evaluation, is a professional duty to reflect on how such evaluative processes provide feedback on teachers' instructional effectiveness. This dimension of accountability must be rooted in professionalism.

This dimension of accountability may differ among accountability approaches as it is likely that teaching professionals may not perceive that they have the authority to make such decisions. Perhaps, in contexts of joint ownership, there may be some confusion about just where does the power to make the evaluative judgment lie. This is point of confusion, perhaps a dilemma that must be resolved within each unique accountability systems. These ideas are presented here as an opportunity to reflect.

In current educational systems, the system level (classroom, school, district, or state) responsible for implementing the policy is deemed the one most capable of

deciding how the policy should be enacted. At the school level, this would involve decisions related to timetabling, the purchase of instructional materials, and organizational procedures. At the district level, this might involve strategic planning or allocation of technical resources. According to Hargreaves (1994) the decision making in restructuring systems is not centralized or decentralized but dispersed among the many level of the system. This is characterized as *responsible decision-making* and reflects the social and cultural elements tied to issues of authority in accountability systems.

Organizational Involvement

This seventh dimension of accountability refers to the depth and breadth of organizational involvement in systemic change. A policy may be the responsibility of one person to implement, either as an individual or as a representative member of a group, but policy implementation often requires the participation of many persons working in a collaborative (vertical) or collegial (horizontal) manner. The former alludes professionals and stakeholders, while the latter to a group of teaching professionals with similar responsibilities. According to Little (1993), a pooling of expertise and experience can build cohesion and promote professional dialogue to examine issues and practices from a range of perspectives and system levels. Such a *professional community* has a positive influence on student performance (Louis & Marks, 1998). These researchers reported that the findings of their mixed method study of 24 restructuring schools identified five elements that capture professional community: shared values, focus on student learning, collaboration, de-privatized practice, and reflective dialogue. The study findings of within school community also observed that school-workplace relationships possessing these characteristics “promote openness, trust, genuine reflection” (Louis &

Marks, p.561). This focus on organizational culture may help avoid what Hargreaves (1992) characterized as contrived collegiality if supportive structures are in place.

Assuming structures that encourage networking, teaming, community partnerships, school-wide improvement planning and regular review of performance contribute to the development of professional community, the depth and breadth of involvement in accountability processes would be extensive. The degree of commitment among school professionals to new policy and change is a function of motivation (Fuhrman, 1999; Leithwood et al., 1999; Sergiovanni, 1996), whether it be in extrinsic form like rewards and sanctions, an intrinsic form such as professional esteem or a combination of both. The question is what fosters such motivation and commitment to change and current research indicates that strategies such as shared decision-making, expanded roles and responsibilities, and professional networking are effective.

The next sub-section examines how the sixth and seventh dimensions of the accountability framework combine to define the organizational context and conditions that shapes the internal accountability system. These dimensions, authority for decision-making and organizational involvement, reflect both structural and socio-cultural elements of the school.

Accountability in Context

Policies do not land in a vacuum (Darling-Hammond, 1990). Policies are apprehended and interpreted by administrators, principals and teachers at all levels of the educational system (Spillane, 1998). What policies and procedures each level of the system is responsible for is determined by the *level of decision-making* and the *organizational involvement* of the broader accountability framework. These two

dimensions determine the structure and the culture of the school context by defining its decision-making authority and its organizational involvement within reform agenda.

The literature on change in schools (Darling-Hammond, 1990; Knapp, 1997; Little, 1993; Louis & Marks, 1998; Peterson et al., 1996) indicates that policies that address only the structure of implementation without considering the needs of the professional community and the process of change have limited success in changing teachers' practices in a meaningful way. Reform agendas are affected by differences in policy interpretation and by differences in how various levels of the organization own the changes and enact them. Horizontal and vertical segmentation can occur when each successive system level deals with policy in a hierarchical manner and no accounting for decisions and actions is made (Spillane, 1998). The same is true for state to district, district to school, and school to classroom relationships. This lack of cohesiveness is credited for differing responses to reform initiatives that have been categorized as *adoption* of policy, *adaptation* of policy and *cooptation* of policy. The challenge of a restructuring agenda is to align the intent of accountability policies at each level of the system so that the aims of the agenda can be achieved at the provincial, district, school, and classroom level. This policy to practice continuum from the provincial level to the classroom level is critical as teachers are the final change agents. Change in teacher practice is nested within the structural and cultural dimensions of the school organization and it is the school environment that is most likely to influence teachers' practice.

Structure and culture are integrated components of school organization. The *organizational structure* refers to the rules, requirements, procedures, and positional duties of the staff. The *school culture* refers to the established patterns of behaviour

(social and professional practices, shared values and expectations, relational trust) that define the professional community. The *mechanisms* for guiding or controlling teacher practices are the formal and informal ways that teachers provide an account of what they do in fulfilling the duties of their position and the same applies for principals. Schools interpret policy based on their professional perspectives and implement it using a variety of processes. Schools are accountable for what they do according to the decision-making authority they possess and the degree to which differing actors (teacher, teams, leaders, principals, administrators, consultants) are involved. How schools enact accountability is an organizational forum, a venue if you prefer, called an *internal accountability system*.

Schools are both structural and socio-cultural entities because the human element operating within the organization plays a role in how rules, routines and procedures are understood and carried out. To appreciate the dimensions of a school-based accountability system, this review now examines what the literature reveals about the constituent components of an internal accountability system.

Internal Accountability Systems

Differing conceptions of accountability systems are discussed in the literature. Several accountability systems use a *top down* perception of accountability called external accountability, while more recent works explore a *bottom up* school-based conception known as internal accountability (Abelmann & Elmore, 1999; Darling-Hammond, 1990; Earl & LeMahieu, 1997; Newman et al., 1997). These two conceptions may be discrete or as newer theoretical directions suggest, they may complement one another. For the purposes of this study, the conception of accountability envisions an

accountability framework that is composed of two accountability systems, one external to the school and the other internal (school-based). The internal accountability system of the school is therefore nested in an external system at the district and state levels. This school-based accountability system is not yet fully understood and is the focus of much attention in educational research today, particularly as it relates to differing external requirements and system-wide approaches to change.

External accountability is designed to promote change in schools and to provide information to the public. As such, it is a key factor influencing the internal accountability system of schools (Abelmann & Elmore, 1999). In this study of internal accountability systems, the internal system is defined as the organizational forum rooted in the context (structural and socio-cultural domains) of the school while the external system is represented by the requirements and consequences that reflect the approach and aims favoured by policymakers. As such, the relationship between the two systems is addressed in terms of the dynamics between component parts of the systems and the degree of compatibility that exists between them.

To develop understanding of the elements of internal accountability, the work of Newman and colleagues (1997) is used to identify four central features of an accountability system: a) information on *school performance*; b) *standards* for measuring achievement; c) *consequences* for success or failure; and d) an influential agent or *constituency who judges (evaluates)* the level of success and distributes the rewards or sanctions. Based on empirical evidence collected from an exploratory study of 20 restructuring schools, these researchers propose that a strong accountability system possesses all four components. Systems that possess standardized test scores, but no standards and no consequences are classified as weak. Mid-range levels of

accountability possess some but not all of the components. In these researchers' model, the source of the consequences is irrelevant and can be externally required, internally generated or be a combination of each.

Newman and fellow researchers (1997) studied schools nested in educational systems possessing strong, inconsistent, and negligible external accountability. From their findings, these researchers conclude that strong accountability can exist *internal* to a school and be independent of district and state systems. In comparing strong *external systems* versus strong *internal systems*, they found that internal accountability was more likely to be associated with high organizational capacity than strong external accountability. Furthering this line of logic, they suggested that since high capacity is associated with quality teaching and student achievement, the promise of internal accountability as a means to influence teacher practice is likely and should be promoted. Evidence of high capacity was not found in schools with strong external accountability systems. In their conclusions, these researchers argue that three issues kept the promise of the popular theory of strong external accountability from working. These issues are summarized as: a) controversy related to standards, incentives, and governance issues, b) insufficient effort to support school capacity with appropriate human, technical and social resources, and c) the failure to recognize the role of internal accountability at the school level.

Abelmann and Elmore (1999) examined internal accountability from the perspective of the individual school 'nested' within differing environments of external policy. Their multiple case study of how schools think about accountability did not examine external systems per se. These researchers explored school site conceptions of teachers' responsibilities and conceptions of accountability embedded in the day-to-day

operations of the schools. These researchers reported that the schools developed normative internal structures that were relatively immune to external influence. They propose that:

“How a school responds to an external accountability system is largely determined, not by the details of those external systems, but by the degree of alignment between the school’s internal accountability mechanisms and the requirements of the external accountability system” (Abelmann & Elmore, p.5).

As the school selection precluded any reference to strong external accountability environments, the relationship between the normative internal structures and the external instruments and strategies remains unexplored.

Based on these descriptive findings, Abelmann and Elmore (1999) propose a somewhat different definition of internal accountability, one based on cultural appreciation for teachers’ perspectives and the role of teacher as change agent within the school community. Their model is comprised of a set of relationships between three factors: conceptions of *responsibility*; *shared expectations*; and internal and external accountability *mechanisms*. The strength of any one factor or the balance between the three characterizes the internal system. Within this school-based cultural perspective, the balance between all three components of the system created different *configurations* that characterize the school’s internal accountability.

Abelmann and Elmore (1999) suggest that school level factors shape a school’s conception of accountability and that external policies may play a role in the process. Their work adds a different dimension to the Newman and colleagues (1997) model, in that it considers the requirements of extant external systems and describes internal

mechanisms that structure the processes through which school-based accountability is enacted.

These and other findings (Cohen, 1995; Darling-Hammond, 1990; Earl & LeMahieu, 1997; Knapp, 1997) identify limitations with external policy-driven systems and highlight the reasons why policymakers should focus on encouraging the type of internal accountability associated with organizational capacity. Findings propose that if accountability is needed, then strong internal accountability complemented with support from informed external agencies in the form of resources, both human and technical, enables change in teachers' practices. The role of external agencies is envisioned as making important substantive contributions to schools "by offering concrete examples of high standards for student performance in specific curriculum areas, approaches to assessment that demand high performance, and reliable ways of evaluating student performance on the assessments" (Newman et al., 1997, p.61). The implication for research is to explore what circumstances can nurture the kind of internal accountability that is associated with high capacity. And furthermore, what mechanisms can be useful for sustaining capacity building and influencing change in teacher practice.

The potential of such rapport between internal systems and external frameworks is debated in the literature. The *top down* perspective of accountability based on social control through legal and bureaucratic authority is challenged in the literature (Cohen, 1995; Darling-Hammond, 1990; Datnow & Castellano, 2001; Fuhrman, 1999; Fullan, 1993; Hargreaves & Fullan, 1992; Hargreaves, 1994; Knapp, 1997; Smith & O'Day, 1991) that highlights the important role of the school and teachers as agents of change. The school reform agenda is witnessing a shift from the 'managed change paradigm' to the 'organizational learning paradigm' as inadequacies of the former to enable real,

sustainable renewal within schools is reported. In his review of the literature on organizational learning, Cousins (1996) proposes that “the installation of apparently highly rational and mechanistic logistic systems coupled with relatively open and free-wheeling interpretive learning systems would appear to fit very nicely within the educational reform agenda” (p.646).

School-level internal accountability systems, as vehicles to effect change, hold promise; and the possibility of state accountability frameworks involving large-scale performance testing to orchestrate a balance between internal and external strategies is receiving much attention (Earl & Torrance, 2000; Fuhrman, 1999; Snider, 2001). Internal and external systems may be discrete or, as newer theoretical directions suggest, they may be complementary under the proper conditions (Earl & LeMahieu, 1997). The components of internal accountability systems are an area that requires further exploration; particularly as such systems are situated to respond to public demand for accountability and to stimulate professional learning.

Reconceptualizing accountability as having two complementary but different purposes is thus proposed in this study. The design of an accountability framework that selects instruments and strategies that balance substantive external direction to guide school practice and ensure transparency with respect for school-based decision-making and professional learning is an innovation worthy of study. How these dual purposes are enabled in real school organizations and community contexts warrants examination, particularly in restructuring systems where the approach to accountability recognizes the role of professional learning to school development (Earl & LeMahieu, 1997; Hargreaves, 1994; Knapp, 1997). The question that this exploratory study hopes to answer with respect to internal accountability is- What is the nature of internal

accountability systems in schools situated in an external accountability environment that recognizes the importance of professional learning in meaningful educational change?

Assessment in Education

The second area of research that informs the propositions of this study is measurement and evaluation. The shift in pre-eminent conceptions of assessment is discussed first. This is followed by a definition of assessment and a list of related assessment terminology used in this study to provide clarification. After this, the characteristics of classroom assessment are explained in terms of the literature that informs this section. A conception of assessment community is also presented to illuminate collective teacher assessment practices. This section concludes with propositions that inform the research questions.

Paradigm Shift in Assessment

In line with the paradigm shift to constructivist theories of teaching and learning, the aims of educational assessment have moved away from a culture of testing to a culture of assessment (Gipps, 1994). According to Biggs (1995), this movement has been framed by changes along three dimensions: a) measurement versus standards, b) quantitative versus qualitative assumptions of evaluating what is learned, and c) de-contextualized learning and testing versus situated learning and assessment. This new assessment is known by several terms such as authentic assessment, performance assessment, standards-based assessment, and formative assessment to name a few. While each term has different connotations, for the purpose of this study the term standards-based assessment will be used because the intent is to capture the focus on

contextualized assessment where student learning is assessed against standards using both quantitative and qualitative dimensions while employing a variety of strategies that capture the process of learning as well as the conceptual knowledge.

The underlying principles and assumptions of educational assessment defined in Gipps' (1994) book, *Beyond Testing: Towards a theory of educational assessment*, are the basis for the list presented here in modified form. Ten principles and assumptions that underlie assessment practice include:

1. Assessing learning is complex and involves examining students, the task, and the context.
2. Assessments use clear performance standards that are shared with students to facilitate self-monitoring.
3. Assessments encourage higher order thinking and are anchored in curriculum content and suitable tasks.
4. Performance tasks are concrete, within students' experience and presented with clear expectations and fair conditions.
5. Assessment criteria are holistic, integrated and supported with exemplars.
6. Teachers are assessment literate and consistent in evaluating achievement.
7. Assessment focuses on the overall quality of the response.
8. Teacher assessment is a key component in teaching because it is interactive, scaffolds learning, and can be evaluated in many contexts.
9. Assessment requires teachers to have subject area knowledge, pedagogical content knowledge and knowledge of students.
10. Assessment is not high-stakes, it assesses against clear standards in a low stakes programme.

These principles present ramifications for schools and for teachers. Classroom assessment requires educators to assess and to communicate about the progress of

students' learning (Stiggins & Bridgeford, 1986; Stiggins & Conklin, 1992) and, in doing so, provide feedback on the intermediate levels of achievement in terms of the skills, the procedures and the progress toward the stated outcomes.

Assessment Defined

With the development of assessment practice, the language of assessment expanded to accommodate the differentiation in purposes and processes. To clarify the meanings of words used in this study, a definition of assessment and related terminology is provided.

Assessment

The classroom assessment *process* possesses a broader definition of what teachers do in the classroom than merely testing. According to Airasian (2000) the term *classroom assessment*, while including traditional paper and pencils tests, encompasses a variety of procedures that involve the “process of collecting, synthesizing and interpreting information to aid classroom decision-making.” (p.293) He differentiates between assessments, tests, measurements and evaluation in the following manner:

1. *Assessment* is a general term that includes all the ways (strategies) that a teacher uses to gather and use information. Tests are one way but other important strategies are observations, oral questions, projects, and portfolios. These other ways have been grouped and called by different names in the literature (performance tasks, authentic assessments, embedded assessments). Assessment can be done for a variety of purposes. If the purpose is *feedback* to inform learning, then it is called *formative assessment*. If the purpose is *grading*, then it is called *summative assessment* and involves a judgment against performance standards.

2. *Tests* are formal procedures for gathering information and are usually paper and pencil instruments requiring students to either select an answer (multiple choice) or provide a response (short answer, essay) format.
3. *Measurement* is the process of quantifying or describing collected data on a performance by scoring it numerically (6/10) or assigning a level descriptor (e.g. 1, 2, 3 or 4).
4. *Evaluation* is the process of making an informed judgment about the quality of a performance. It occurs after assessment information has been gathered, synthesized and reflected upon. Evaluation is the process of making a judgment based on a comparison of student overall performance against pre-established performance standards. (Airasian, 2000, pp 9-12)

From Airasian's (2000) perspective, *grading* involves a teacher judgement based on a comparison of individual student performance (*assessments*) against set performance standards (*evaluative criteria*). Popham (1999) also distinguishes grading from evaluation using a decision-making perspective. He differentiates between decisions using evaluative criteria that lead to assigning a grade versus those that lead to conclusions about the quality of the instructional program. He defines *grading* as a function of student performance appraisal and uses the concept of evaluation to refer to program appraisal, in other words, *program evaluation*. He acknowledges that program evaluation is indirectly teacher evaluation because it involves a reflection on the quality of the instructional program. Linn and Grunland (1995) distinguish between assessment and evaluation by describing *evaluation* as the step following assessment where the judgement of the information leads to a *decision* that is actioned.

For the purposes of clarity in this study, student grading will refer to the result of a decision based on summative assessment judgements and evaluation will refer to appraisals of teachers instructional programs. This latter term will also be used when

addressing issues of school wide appraisal of assessment results from external or internal performance information.

Within the context of this study, the use of the term summative assessment is qualified for purposes of clarity. When a grade for mathematics is condensed to one achievement report (i.e., A, B or C, etc.), this requires the teacher to action a decision based on an overall evaluative judgment about a student's performance across each of the mathematical strands in the curriculum. When reporting policy further requires that each strand be graded individually, there may be an expectation that each strand's final grade reflect a decision based on the review of all summative assessments for individual topics within the strand. A final grade therefore represents a decision that is actioned based on an overall evaluative judgement of several summative assessments.

Classroom Assessment

Classroom assessment has received much attention since Stiggins and Conklin (1992) developed a comprehensive questionnaire for the analysis of classroom assessment procedures. Purposes for classroom assessment (Gipps, 1994; Shepard, 1995; Stiggins et al., 1992), methods of assessment (Black & Wiliam, 1998; Stiggins & Conklin, 1992), recommendations and models for classroom assessment (Costa & Kallick, 1995; Earl & Cousins, 1996; Kulm, 1994; Harlen, 1994; Prestine & McGreal, 1997; Schafer, 1991), criteria for assessment selection (Popham, 1995; Ryan, 1997), assessment for decision-making (Costa & Kallick, 1995; Wilson, 1990), and assessment feedback (Black & Wiliam, 1998; Natriello, 1987) are areas that have been addressed in the literature to provide background for what classroom practice should involve.

For the purposes associated with this study, the characteristics of classroom performance assessment are organized along the following dimensions: a) instructional purposes, b) criteria for selection, c) methods of assessment, and d) feedback procedures. (see Figure H3) The quality of teacher assessment is not addressed in this study, being outside the scope of this work.

Instructional Purposes

Teachers assess students for a variety of instructional purposes in the new paradigm of classroom assessment. The purposes are grouped under the categories of diagnostic, formative, summative, student management, test-taking preparation, and accountability.

Diagnostic assessment is done when specific information is needed to locate difficulties and determine potential for special needs students. It is less structured when it takes the form of classroom pre-assessment and is used to determine strengths and weaknesses among students. Pre-assessment is frequently used to form instructional groupings or to plan instructional units (Stiggins & Conklin., 1992).

Formative assessment provides information that is used by the teacher and the student to modify their work. It involves feedback that clarifies expectations, recognizes improvements and indicates next steps. It may also address affective issues and may be a form of praise and motivation. Formative assessment may also provide feedback on instructional effectiveness and program quality (Stiggins et al., 1992).

Summative assessment also involves gathering information about student learning but it differs from formative assessment because the information gathered is used to grade students for administrative purposes such as promotion and certification.

The grades are symbolic representations of an evaluative judgement (O'Connor, 1999; Popham, 1999).

Student management is another instructional purpose for assessment and happens when teachers need to motivate students or guide behaviour. It can be linked with other purposes of assessment or may be a form of classroom management. (Stiggins & Conklin, 1992).

Test-taking preparation is another purpose for assessment. Its function is to expose student to test-like situations. The assessments model the test-taking format and strategies (Stiggins & Conklin, 1992).

Assessment for the reasons of *accountability* is another purpose that is a common practice. Student performance may be assessed for certification purposes (individual accountability), for programming feedback (school accountability), or for community awareness (community accountability). The intent depends on the accountability approach favoured by educational stakeholders and the manner in which the assessment results are reported (Leithwood & Earl, 2000).

Criteria for Selection

In classrooms, teachers select and use assessments for a variety of reasons. Stiggins and Conklin (1992) identified the following criteria that teachers may use when selecting or designing assessment tasks.

Methods reflect intended outcome. The methods of the assessment tool or strategy match the materials used in the instruction and the learning outcomes taught. The learning target is the focus in the assessment selected. The assessment is useful for measuring the knowledge and skills taught (content knowledge, procedural knowledge, application in problem-solving, or communicating reasoning). An example would be the

use of journaling in mathematics to assess students' reasoning and ability to communicate mathematical thinking.

Adequate number of questions to measure performance. The assessment tool or strategy possesses an adequate number of items to clearly demonstrate mastery of the learning outcomes both in frequency (number) and comprehensiveness (skill level).

Interference controlled. Distracters to fair assessment are minimized by assessment characteristics such as increasing objectivity (clear criteria), control of cheating (assessment design) and availability of time.

Methods fit the purpose. The assessment methods are suited to the purpose. If the assessment was designed to be formative, the results are used to assist with instructional planning or scaffolding learning but not, for example, for a summative grade.

Economy of use. This category refers to the time and effort needed to prepare, conduct and record the assessment. This may also be a factor related to the origin of the assessment since assessments embedded in programs are credited with efficiency as well as suitability.

Methods of Assessment

Drawing upon several sources (Earl & Cousins, 1996; Ontario Ministry of Education, 1999; Stiggins & Conklin, 1992), assessment methods can be grouped in three ways: a) recitations, b) paper and pencil instruments, and c) performance tasks. Several methods are itemized here but the list is not all-inclusive.

Teacher developed paper and pencil assessments include tests, quizzes and activities designed by the teacher. There is concern that these tests lack rigor and reflect what was taught as opposed to the expectations in the curriculum.

Text-embedded tests and quizzes include the tasks designed for review and summative grading: including cumulative activities found within the textbook or teacher guide materials. They match both the material taught and the scope and sequence of the program.

Standardized tests are psychometric tests designed to evaluate student potential or track progress in basic skills over time. They are usually used for diagnostic and remedial purposes. These tests are designed by measurement experts.

Classroom recitation refers to a variety of communication based on approaches such as: a) question and answer, b) interviews, c) conversations, and d) comments from other observers.

Performance based tasks require strategies for gathering observations of the thinking, problem-solving skills and application of concepts and terminology to communicate learning and achievement. Tools such as checklists, anecdotal records, appraisal scales to assess process or products, and rubrics guide the assessment of the performance tasks. A collection of products (portfolio) that model the processes and products reflect criteria and demonstrate learning is another tool. A performance-based assessment is characterised by a formal observation and judgement of a student's skill in carrying out an activity or producing a product (Joint Committee on Standards for Educational Evaluation [JCSEE], 2003). In the arts, performance assessments require professional judgement based on expertise and knowledge of the field.

A use of a variety of assessments methods is encouraged as this facilitates higher order thinking, shows awareness of different learning styles, reflects respect for cultural differences and addresses the issue of test anxiety (Earl & Cousins, 1996).

Feedback Procedures

Feedback is an integral part of classroom assessment and may include oral, non-verbal or written forms. Feedback is also a critical component of the teaching, learning and assessment spiral (Costa & Kallick, 1995). As such, assessment is a *moment in learning* because students are active in reflecting on their understanding and misunderstandings, and they are more aware and motivated to resolve the discontinuity. The nature of feedback in classroom assessment is a key feature of standards-based assessment and its formative function. Black & Wiliam's (1998) review highlights four elements in a feedback system: a) data, b) a reference to expectation, c) a means to compare the different levels, and d) a strategy by which input can reduce the gap. The following feedback procedures contrast different strategies and inform on their effectiveness.

Comments versus symbols. Findings from the literature (Black & Wiliam, 1998; Sadler, 1987; Stiggins & Conklin, 1992) suggest that verbal praise and supportive comments motivate students, but have modest effect on performance. Students who receive written comments that are task oriented tend to be more committed to the learning process than students who receive praise, grades or no feedback at all. Task commitment is associated with improved performance by students and linked to higher achievement.

Guidance versus praise. This aspect refers to the nature of the feedback and whether it draws attention to the *task* versus directing attention to the *self* (the learner), as in rewards or sanctions. Attention to the self and away from the task favours a negative effect on student achievement. Black and Wiliam (1998) note that effective teachers tend to praise less frequently and provide more constructive feedback.

Germane versus irrelevant. Feedback should be task specific and relate to the criteria and expectations that the learning was intended to address.

Immediate versus delayed. The immediacy of feedback is linked to the assumptions underlying learning theory. Frequency is useful to *scaffold* present understanding to new understanding. The benefit of frequency can be impacted by the type of feedback and its quality.

Product versus process. Feedback that focuses on *the learning process* promotes greater achievement by students than feedback that focuses on the end *product (achievement)*. The former is situated in understanding *how* to solve the problem, while the latter recognizes that a problem is solved (Black & Wiliam, 1998). The learning process orientation supports higher motivation and achievement outcomes for most students, when compared to the product orientation.

Self and peer assessment. Self-monitoring was found to be an inescapable feature of formative assessment in a number of studies reviewed by Black and Wiliam (1998). Self-evaluation is intrinsic to reflecting on one's work. Through feedback on the *self-assessment process* itself, students learn to better approximate their perception of the quality of their work to correspond with the judgements of others. *Assessment conversation* is identified as a means to encourage this skill and can help explain about criteria, what needs to be improved and how to go about it. *Peer assessment* is developed through the same process but is less successful in contexts where groups assess groups.

Assessment practice has become increasingly complex as the purposes, methods, criteria and nature of feedback are informed by current research employing constructivist learning theory. Assessment literacy is a well documented concern. This

emerging dilemma for increased capacity building has focussed attention on supportive collegial contexts and school-based professional learning to address these teacher development issues.

Assessment Community

Just as the characteristics of classroom assessment are qualified in the literature, studies of school assessment practices (Hall & Harding, 2002; Monson, 1998) provide insight into school level characteristics associated with modern principles of assessment. These school level dynamics are examined here using the dimensions identified in Hall and Harding's (2002) empirical work on teachers' interactive assessment practice and represent an interpretation of their work. These researchers used qualitative analysis (Miles & Huberman, 1984, 1994) to examine the school-level mechanisms and structures that promoted sharing of assessment criteria among the staff, students and parents of six English schools using level descriptors for grading student achievement. The five elements they identified include: a) goals; b) assessment instruments; c) assessment methods; d) personnel; and e) values. These five dimensions build an understanding of school assessment dynamics and contrast teachers' response as either an *individualistic* enterprise or a *community* endeavour. Figure H4 provides a graphic version of the information. This dichotomy reflects the same paradigm shift found in assessment (Biggs, 1995; Gipps, 1994) but frames the paradigm shift from a professional practice perspective. In other words, as the conception of assessment has changed, so has the conception of teachers' assessment practices at the individual and school level.

Goals

The goals of an assessment community reflect a *compliance* and *acceptance* of policy initiatives versus the *reluctant compliance* and *resistance* that is termed individualistic assessment (Hall & Harding, 2002). The willingness to use standards based assessment practices and to invest professionally in learning and acquiring the needed knowledge and skills is evident. An assessment community is further characterized by its collaborative nature.

Instruments

The use of diverse assessment instruments, strategies and tools for assessing and recording student learning is anticipated in assessment communities. In standards-based assessment there is application of *consistent standards* through key instruments. Rubrics or achievement charts (level descriptors) help teachers evaluate student achievement against a standard (standards differentiate the criteria for each level of achievement along a ranking scale determined by state-level agencies). In contrast to an assessment community, little or no sharing of what standards mean is termed individualistic. Active use of portfolio to collect samples of student work for evidence of growth and development over time is another feature of an assessment community. The focus is on assessing student learning as well as recognizing the end result. Finally, Hall and Harding (2002) explain that in assessment communities, teachers own and use exemplars to inform teaching, learning and assessment. These may be a mixture of school and system materials but are not usually commercial products.

Methods

Assessment methods and procedures tend to diversify in assessment communities and teachers use multiple sources in both the collection of information on student

achievement and the assessment of performance levels (Hall & Harding, 2002). An assessment community plans the *collection of assessments* that will be used in the process of instruction. A variety of forms (strategies and tools) are included and the assessment process is embedded within the instruction as much as possible. The assessment focuses on process goals (how to) as opposed to a summative perspective intent on product goals alone (outcome). Furthermore, assessment communities know and use assessment terminology correctly to facilitate clear communication about goals, processes and results with students, parents and colleagues. They often refer to moderation as a team-based approach to assessment that involves collaborative development of assessment tasks and tools and active discussion about evaluative criteria with the intent to develop consistent assessment processes and judgements of student work.

Personnel

The depth and breadth of persons involved in the assessment process is expanded in assessment communities (Hall & Harding, 2002). School improvement planning is a collective effort by principal and teachers. It is a characteristic of assessment community that has several discrete aspects which may include: regular review of school, district and state level testing results; identification of program strengths and weaknesses; shared decision-making on key issues for school improvement and on how changes will be integrated; teacher ownership for professional learning opportunities and sharing of expertise; school commitment to reviewing district planning documents and establishing use within the school; and consultation with parents and the local community to establish action plans. Teachers and parents jointly review assessment and testing results and share in the development of individual learning plans for students.

Values

The value system of an assessment community differs from that found in contexts where individuals operate independently (Hall & Harding, 2002). Assessment communities value the collaborative assessment process, understand the role of assessment in improving instructional practice and articulate the benefits of student self-assessment and learning skills development. Teachers and principals discuss what resources (human, technical, financial) are needed and seek support from the district level (expertise) and the community level (parent councils, community organizations).

Assessment communities are collaborative venues to support quality instructional programs and focus the aim of schooling on the promotion of student achievement. They represent a combination of agreed upon goals, tools and procedures, personnel involvement, and values (Hall & Harding, 2002). They are more than the sum of the individual teachers and are an appropriate aim for policymakers and educators at all levels of the education system.

The benefits of assessment (Black & Wiliam, 1998; Smith & O'Day, 1991) to student learning and instructional practice are well documented. The positive outcomes include: a) clear academic standards for evaluation and reporting, b) recognition of *process* as well as *product* in student assessment, c) teacher and student ownership of assessment, and d) the promotion of formative assessment. Two perspectives on how standards-based policy should be implemented are held. Cross & Joftus (1997) contend that standards-based assessment practice fosters program reviews and instructional improvement when increased incentives, training, and resources are provided to ensure effective implementation. However, other researchers (Fullan et al., 1991; Hargreaves, 1994; Darling-Hammond, 1990; Peterson et al., 1996) caution that more than effective

training is needed. This second perspective advocates for a professional learning approach that situates teacher learning in school-based contexts where collegiality and professional esteem are the critical factors that motivate. However, empirical works (Borko, Mayfield, Marion, Flexer, & Cumbo, 1997; Cizek et al., 1996; Daniel & King, 1998; Stiggins and Conklin, 1986) on the influence of assessment to inform instructional decision-making do report that the effectiveness of assessment to inform instruction is limited by teachers' limited understanding of classroom assessment purposes, methods, criteria, and the role of assessment feedback within the teacher-learning dynamic. This knowledge constitutes assessment literacy. The literature flags that successful implementation of standards-based policies must consider three influences: motivation to change, meaningful opportunities to learn, and the availability of quality resources.

In schools where system restructuring has been underway for almost a decade and large scale performance testing instituted, what type of teacher assessment practice can be expected and how well have the characteristics of a viable assessment community been established? This study investigates such contexts.

Theories of System Change

The third area to be explored for the purpose of this study has to do with theories of system change. In considering system-wide change, the word *reform* implies a different set of assumptions than does the word *restructuring* and this is elaborated in this section. The assumptions held by policy makers and stakeholders responsible for change, whether through reform or restructuring, underlie the aims and objectives of new policies, particularly as it applies to issues of accountability. The intent of this

section is to review background issues and perspectives related to system change and to accountability agendas associated with those changes. This part of the literature review differentiates the philosophical perspective that informs a model of system change developed within the Province of Ontario, Canada, specifically as it pertains to the accountability system in place. The operant external accountability system is described in order to inform on the nested context under which the internal accountability systems examined in the school case studies are presented. Propositions related to the Ontario accountability approach complete this section, after which the conceptual framework is outlined.

Background Issues

System change is described as *coherent policy*, as *alignment of system elements* and as *standards-based reform*. These three terms have a similar purpose in mind- an expectations that a degree of compatibility exist between the expectations of the educational system at the policymaking level and the implementation of that policy at the district and school levels. Bringing compatibility of processes and shared expectations is one intention of the standards movement guiding educational accountability programs in Ontario. Associated with this movement are two, somewhat different perspectives of system change known as *educational reform* versus *educational restructuring*. Hargreaves (1994) distinguishes restructuring from its antecedent of educational reform by clarifying that “Change by *reform* sought to mandate improvement upon teachers by bureaucratic control and compliance, rather than by supporting teachers in improving themselves and creating restructured opportunities for them to exercise their professionalism”(p. 52). The value in considering the differences

between these two perspectives comes from research in educational change and professional learning.

The literature on educational change (Fullan, 1993; Hargreaves, 1994; Harris, 2000; Hibbard, 1996; Kruse, 2001) reports the effectiveness of school-based structures and interactive processes to address complacency and trigger teacher ownership for school improvement. Such findings support Little's (1993) six principles associated with professional learning and sustained change in teacher practice as: a) meaningful intellectual, social and emotional engagement with ideas, materials and colleagues, b) situated learning, c) respect for informed dissent, d) awareness of nested school contexts, e) use of inquiry techniques and teacher expertise, and f) assurance of bureaucratic restraint and balance between the interests of the individual teacher and the school institution. Further support for these principles is found in more recent studies where meaningful change in teacher practice was limited because of compliance orientations (Datnow & Castellano, 2001) and high-stakes contexts (Mintrop, MacLellan & Quintow, 2001). School structures that provide for collaborative venues enable teacher dialogue and facilitate professional learning that focuses on issues of teaching and learning (Kruse, 2001; Peterson et al., 1996). These findings highlight the importance of *situated learning* and *school support systems* to change in teachers' practice. Restructuring is therefore based on professional learning situated and supported in schools.

The importance of accountability policies and instruments to guide the direction of school change and to provide opportunities for professional development at the school is pivotal to effective system restructuring (Darling-Hammond, 1990). Hargreaves (1996) proposes that collaborative school structures that integrate *knowledge utilization* and teachers' *professional practical knowledge* are the means to inform and motivate

school improvement, but little is known about such structures, their mechanisms, or how they function in restructuring contexts. Current accountability agendas are increasingly based on large-scale performance tests that policy makers envision as making substantive contributions to school decision-making and professional learning through provision of external measures of school performance while simultaneously ensuring transparent accountability. Concerns reported with such policies include *how* results are used and in *what ways* the tests impact teachers' practices (Elmore, 1995; Firestone et al., 1998; Herman, 1997; Linn, 2000). These issues need to be explored because school principals and teachers interpret and implement policy through: a) talking about it, b) processing the meaning, c) inquiring into it, and d) reformulating it taking school circumstances and the students they know into account (Hargreaves, 1996). Therefore, these educators are ultimately *responsible* for implementing the policy within the school context and in the process likely to *adapt* it.

Within the restructuring movement, Hargreaves (1994) further describes a fundamental contrast between *restructuring as bureaucratic control* and *restructuring as professional empowerment*. The contrast between these two positions is a belief, or perhaps, a practice about where authority for education should lie. This means that tensions between the hierarchical-corporate perspective and the democratic-cultural views are problematic. Hargreaves contrasted the fundamental dilemmas as a) vision versus voice, b) mandates versus menus, c) trust-in-persons versus trust-in-process, and d) structure versus culture. Much debate remains (Fuhrman, 1999) on where authority should lie and what new forms of accountability best serve.

The nature of restructuring is reflected in the parameters of decentralization, empowerment and client service. As Gurney and Andrews (2000) explain, the term is

used within the “connotation that decisions must be decentralized to the level at which the problems themselves actually occur” and imply that educators must be empowered to frame problems and propose solutions. The permutations and combinations of restructuring formats within three tiered organizations are diverse and this illustrates how the breadth and depth of stakeholder involvement requires extensive collaboration and cooperation. Restructuring systems are challenged to find the appropriate balance between top down oversight and bottom-up motivation for meaningful institutional development (Fullan, 1998; Hargreaves, 1994; Hatch, 1998; Leithwood & Earl, 2000). The implementation of compatible strategies that aim to balance the need for public accountability with local ownership for school improvement is underway in Ontario (Earl & Torrance, 2000; Snider, 2001), but evidence of how such aims are realized and the impact on instructional effectiveness is limited.

A conception of accountability is shaped by public beliefs and by political will. In post-modern society, stakeholders must question if one view of restructuring will satisfy the demands and needs of a complex, nested social institution like the public education system. Policymakers and researchers must ask if a middle ground position might offer the benefits of both perspectives. The restructuring initiative in the province of Ontario provides an opportunity to examine one such accountability framework.

The Change Process in Ontario

School improvement and quality education for all students have been the impetus for the systemic change movements of the past two decades. In 1993, the Government of Ontario established The Royal Commission on Learning to conduct extensive public consultation with educators, policy makers, parents, students and interested citizens

across the province to discuss concerns about the state of education and to address issues of accountability. The Commission's final report contained 167 recommendations, clearly identifying "accountability as a burning and legitimate issue for many Ontario parents and citizens" (Green, 1998, p. 8).

The Education Quality and Accountability Office (EQAO) was created in 1995 in response to public demand for greater accountability. Its mandate, as an independent arm's length agency, is to conduct province-wide student assessments, coordinate participation in national and international tests, and develop an Education Quality Indicators Program (EQUIP). Its responsibilities are to report to parents and the public about student achievement and educational quality in Ontario and to make recommendations for school and system improvement (EQAO, 2001-2002). Its use of large-scale assessments to monitor student achievement and provide feedback on school performance was part of a set of coherent change policies initiated as a result of the Royal Commission Report.

Other accountability policies associated with the comprehensive systemic reform agenda arising from the Commission report, *For the Love of Learning* (1994), include:

1. A new Ontario Curriculum that listed specific expectations and a standard of achievement for each grade level (1997).
2. A standardized report card that aligned with the curricula and the achievement levels (1997).
3. A secondary school reform program (1999).
4. A professional College of Teachers with authority to license, govern and regulate the practice of teaching (1996).
5. A program to establish school councils to act as advisory bodies at the school and district level (1996).

6. A new student-targeted approach to funding to increase the accountability of school boards (1998).

The intention of the Royal Commission was to create an open, transparent, collaborative and accountable educational system. This study examines one aspect of the restructuring agenda, the accountability program and its large-scale assessments, in order to describe how external accountability systems interact with school internal accountability systems and to describe the mediating influence of internal accountability on teachers' assessment practice. This section now examines Ontario's accountability framework and situates it within a continuum of approaches to accountability.

Ontario's Innovative Accountability Framework

Earl and Torrance (2000) report on the approach to accountability that has been developed in the province of Ontario. It is an emergent conception of accountability that extends the ideas presented in Fuhrman's (1999) tentative model of *new accountability* because it employs a conception of professional responsibility in a somewhat different way.

In an attempt to situate the Ontario approach to accountability within a continuum of accountability frameworks, the classification schema and theoretical principles of theorists (Adam & Kirst, 1999; Hargreaves, 1996; Leithwood et al., 1999) are interpreted and organized into chart form for the purposes of this study (see Figure H5). The chart tentatively captures the key dimensions of different approaches to accountability as defined in the literature, including the emerging professionalism model.

In contrast to a bureaucratic approach, where large-scale accountability testing is a means to induce change through the application of extrinsic rewards and sanctions (Fuhrman, 1999), the Ontario approach appears to be an initiative with a much different role. Ontario's approach is one that perceives large-scale assessments as a means to stimulate conversation about school performance, collective expectations and school improvement planning. This approach to accountability and change is best qualified as restructuring based on professionalism.

The bureaucratic approach to accountability is well established and has been used in a variety of restructuring contexts with varying degrees of effectiveness. The professionalism approach is an emerging phenomenon and its effectiveness to foster and sustain change through large-scale assessment programs designed to promote systemic reform remains inconclusive at this time. Early studies (Earl & Torrance, 2000; Snider, 2001) report progress in promoting capacity building and change in teachers' practice but modest progress in stimulating assessment conversations about school improvement and student learning beyond school personnel. Questions remain about the time needed to implement such ambitious systemic change, the compatibility of the accountability instruments, the sustainability of school development processes, and the appropriate balance of guidance versus control across the levels of the system.

The notion that reform policies can be promoted by an interactive dynamic which addresses coherence in the system through a centralized coordination of curriculum, standards and assessment is still to be determined (Smith & O'Day, 1991). The findings of current works have not affirmed this position is workable but the relationship between policy directives and school practice may be observed through what Earl and LeMahieu (1997) describe as *assessment as conversation*. Assessment as

conversation refers to discussions of accountability assessments that involve teachers, administrators and parents in discussions about expectations, performance and instructional planning. Does this new definition of assessment link the aims of accountability with the aims of learning? If so, can the interaction between external and internal systems be like a conversation that *balances* the need for coherence in standards and expectations with the needs of informing and enhancing teacher practice? This *assessment as conversation* is a concept and a process that is explored and described in this study.

Embedding Accountability and Improvement in Assessment

Earl and Torrance (2000) articulated the aim of the EQAO to develop a quality large-scale assessment program in such a manner that it would also change teacher classroom practices in positive and proactive ways. In order to avoid the disadvantages associated with many large-scale (often high stakes) tests, these researchers described the following issues as important considerations in the EQAO planning process:

1. Assessment is a powerful lever for change and assessment reform can be designed to exert external pressure for compliance with mandates or it can be envisioned as an internal process building on teacher learning and professionalism.
2. The purposes for assessment determine the approach and the design of the assessment program; political as well as educational agendas must be appreciated.
3. Large scale assessments can become disconnected from learning if it is used to reward or punish teachers and schools.
4. The shortcomings of multiple choice tests to elicit higher order thinking must be considered as well as the complexity and difficulties inherent in performance based tests.

5. Anticipating negative effects must be factored into the process because teaching to the test impacts on schooling practices and unanticipated moral and political issues unrelated to testing issues frequently occur.
6. Research on school effectiveness indicated that while schools do make a difference on student achievement, contextual information is needed to determine other influences.
7. Accountability informed by large scales assessment requires scores that can be reported in transparent ways understandable by students, parents and the public.

In designing a program to address the dual aims of accountability and improvement, the EQAO employed a framework based on Fullan's (1993) change theory and on a definition of accountability as a process that ensures schools and school systems meet their goals through a dialogue among stakeholders.

According to Earl and Torrance (2000), the assessment program format developed from this commitment possessed the following characteristics: integrated performance assessments tasks (Arter, 1999) that called for complex thinking and the ability to apply knowledge; teachers trained to implement the assessment; questionnaires that gathered information from teachers, administrators, parents and students about instructional programs, resources, student backgrounds, etc.; holistic scoring using a four-point scale in line with curriculum expectations and the criteria set in the provincial achievement charts; teachers trained to score the assessments using moderation processes; individual student, classroom and school reports prepared to facilitate discussions and improvement planning; EQAO provincial reports signalling appropriate and inappropriate interpretations of the recommendations in all reports and media coverage; provincial, district and school reports designed to be simple, accessible and transparent in order to motivate and initiate conversation among and between all

stakeholders; and a public communication strategy focussing on improvement and learning.

From the initial results of the assessment processes during its first two years, the EQAO *Provincial Report on Achievement* (1997) identified 21 recommendations specific to the primary and junior grades related to the reading, writing and mathematics programs. Four general recommendations for schools that arose indicated the need to create a culture of assessment, the need to focus on the curriculum, the need to support student learning, and the need to communicate and interpret assessment findings (p.64).

Early findings based on the 1998 survey of elementary schools by Earl and Torrance (2000) revealed that the impact of the Grade 3 provincial assessments was promising but not without limitations. Approximately 840 schools responded to questions about participation in the assessment, use of the results, changes in the school instruction, implementation of recommendations. From survey results that the authors caution may overestimate the impact of the assessment, the findings indicated that many teachers and principals were using the assessments to focus attention on curriculum, instruction, assessment, and resources. However, they also revealed that EQAO recommendations to enhance face-to-face discussions about assessment processes and results were very limited. Some sharing of school results happened at parent meetings and staff meetings. Teacher-teacher and teacher-principal conversations were common practice but school-community and school-parents assessment conversations were not. Reports of school results were often incorporated within routine newsletters. Their findings reported that students were not involved in self-assessment and evaluation. The study did not provide explanation for this situation but reference to other, earlier

research works was used to speculate on the reasons and these included assessment illiteracy and teacher vulnerability.

Another study (Snider, 2001) on the impact of structural and cultural conditions in a sample of Ontario schools was done to learn whether the Ontario's accountability approach resulted in gains in school aggregate scores in Grade 3 writing. Findings based on years one and two of the EQAO testing suggested that school results improved with familiarity with the test, but did not show a correlation with change in instructional practices in high performing schools. Trends in the study data suggested that high performance schools benefited most from familiarity with the test, professional development on interpreting scores, and explaining results to parents. The latter finding in this mixed methods research appears to contradict the Earl and Torrance (2000) conclusion that, in general, teachers were not engaging with parents and the community in meaningful ways. At issue here may be the nature of communication with parents, where the former study views sharing student and school results with parents as communication, while the second sought to determine if EQAO results were discussed with parents and recommendations to involve school and community stakeholders in supporting school programs and developing educational plans for students. Snider (2001) also found no correlation between teacher collaboration and student achievement at this early point in the implementation of the EQAO testing program.

These studies were conducted in the stages of the EQAO testing program and focussed on the Grade 3 assessments. Since that time the design of the provincial tests has been refined and the testing has been extended to include Grade 6 reading, writing and mathematics at the elementary levels, Grade 9 mathematics, and Grade 10 language literacy. Furthermore, the impact of the EQAO recommendations and district / school

improvement plans requirements were preliminary. Access to the standards-based assessment resources such as EQAO anchors, exemplars, and performance-based instructional materials were not available. Neither study appears to address the nature of internal accountability systems or the influence of the accountability mechanisms to facilitate the complexity changes within school organizational structures and interactions.

With the EQAO assessment program established for six years and Grade 6 provincial testing active for four years, a study of the characteristics of school internal accountability systems and their relationship to the innovative Ontario accountability program was warranted. The research questions that guided this case study research are presented next, followed by the conceptual framework that situates this exploration of school-based accountability and operationally defines the conceptual model of internal accountability systems used in this study.

Research Questions

This study is designed to examine and describe the extant internal accountability systems of two schools nested in the same school district in the province of Ontario. A look at the Ontario accountability system is warranted due to its relatively new and innovative approach to system change and the potential it may have to stimulate the kind of school internal accountability system that research in the field states credits. The study is intended to reveal issues of policy implementation, taking into account the perspectives of teachers and the complex dynamics found in school organizations. The mechanisms that shape the school internal accountability systems are identified and the decisions, actions and perspectives of the members of the professional communities are

revealed. Since internal components are connected to those of external systems, the relationship between internal and external systems is explored.

This study also addresses propositions rooted in a review of the literature that ties the dual purposes of accountability as assessment, to improve teachers' practice and to make school accountable for quality instruction (student outcomes). Empirical evidence is needed to reveal if these dual purposes are achievable. Study results may provide suppositions regarding the effectiveness of the current accountability agenda to maintain a balance between the requirement for compatible policy and the flexibility to address context specific needs.

Another intention of the study is to describe current teacher assessment practices at each school and provide insights into the potential of internal systems to influence school-wide response by fostering change in teachers' assessment practices. This study anticipates different school-level responses to external policy and different configurations among the elements that make up the internal accountability systems in the two school case studies. Based on these intentions, the following research questions guide this investigation:

1. What is the nature of school internal accountability systems in a context of external accountability?
 - a) What is the strength of the extant internal accountability systems?
 - b) What factors impact on these systems?
 - c) What is the relationship between the internal and external accountability systems?
2. What is the mediating influence of internal accountability, if any, on teachers' assessment practices?

Schools are complex entities and capturing how teachers perceive and respond to policy initiatives requires close examination. A collective case study of schools is a means to better understand the dynamics, perspectives and responses of schools in restructuring contexts. The study of schools within a common district and provincial context invites cross-case analysis of the internal accountability systems and the influence such forums may have on teachers' practices.

The Conceptual Framework

This study of internal accountability systems is situated in a provincial context where system-wide restructuring is underway. Policy-to-practice connections are problematic and careful analysis is needed to reveal the components of the internal accountability systems, the factors that influence these systems, and the nature of the relationship between the external framework and the internal response, particularly as it relates to teachers' assessment practice. The analytical framework is stated first and the conceptual model used in this study is presented afterward.

Framework for Analyzing Systemic Change

The problem of describing and analysing policy-to-practice connections in restructuring systems is complex because, as Knapp (1997) identified, no single theoretical perspective can address all the outstanding issues. In his view, these connections are best understood using an analytical framework informed by three fields of study: a) policy implementation, b) innovation and change, and c) professional and organizational learning (capacity building).

Policy *implementation*, once viewed as a matter of simply defining the policy, setting mandates, providing the guiding frameworks and checking on the results, has recognized limitations. A focus on compliance only did not bring the desired sustainable change in teaching and learning at the classroom level (Elmore, 1990; Fullan, 1993; McLaughlin, 1987). Research (Chrispeels, 1997; Cohen & Ball, 1990; Meyer, 2002; Spillane, 1998; Spillane & Thompson, 1997) discusses the way local contexts interpret policies differently and influence policy implementation. According to Knapp (1997), school reforms that were unsustainable or inconsistent were found to fail because decision-makers did not fully appreciate teacher-related constructs such as “capacity, will and mutual adaptation” that operate at schools over time.

Studies of *innovation and change* highlight the processes by which resistance to change can be overcome (Fullan & Stiegelbauer, 1991; Hatch, 1998; McLaughlin, 1987). The role of active change agents who promote awareness and knowledge of the new policy and facilitate its use through school-based strategies are credited with addressing teachers concerns. New definitions of innovation envision change as a process of institutional development that increases school capacity and performance for continuous improvement (Fullan, 1993). This theoretical perspective identified distinct stages in the process of acceptance of change. However, according to Knapp (1997), this perspective alone limits appreciation for the personal construction of teachers’ knowledge, philosophy and beliefs about teaching and learning.

Studies situated within a professional and organizational learning perspective (Chrispeels, 1997; Cohen, 1995; Darling-Hammond, 1990; Knapp, 1997; Spillane, 1998) reveal that the goals of system change are ambitious; and that influencing change in teacher practice is problematic. Examinations of activity at the school and district

levels have found the following constraints: a) policy must be better communicated to teachers and administrators, b) policies do not land in a contextual vacuum at the school and district level, c) teachers teach from what they know and understand, and d) the process of change is slow and difficult (Darling-Hammond, 1990). This perspective addresses the importance of learning but doesn't explain issues of motivation and sustainability over time nor does it address how to put learning into practice.

Knapp (1997) recommends that qualitative methods are needed to capture the strategic limits of systemic reform policies, their instruments, and strategies to both guide and monitor what is happening in school (teachers) practices. His analytical framework for examining multi-faceted *policy-to-practice* relationships addresses the three theoretical perspectives in the identification of four essential elements: a) the target of the policy, b) the policy instrument, c) the avenues of influences, and d) the context and conditions that mediate policy influence. These elements of systemic reform thinking are useful for interpreting the ways that accountability policy is enacted at the school level because this enables analysis of the specific aims, instruments, tools and strategies that make up the internal accountability system of the school. It also highlights the avenues of influence by which the ideas of teachers may be influenced.

Using an adaptation of Knapp's analytical framework, this study looks at how change messages are communicated to schools and examines what specific accountability instruments, strategies and tools are in place to ensure both accountability and enhanced teacher practice.

Ontario's emergent professional *approach* to accountability is described in the literature. Together with the stated dual aims of the EQAO accountability program, a conception of external accountability systems is formed. For the purpose of this study

one more element is added to Knapp's original analytical approach, the aims of the policy, as this extends the analysis to account for all dimensions of accountability arising within the literature review.

For the purposes of this study then, the systemic reform *policy* is standards-based achievement, the policy *aims* are accountability for improvement and teacher learning. The *target* of the accountability policy is teachers' practices and the *policy instrument* is the EQAO assessment program (EQAO tests, standardized report card, provincial achievement chart, public reports, etc.). The *avenues of influence* are the conduits that influence decision-making (i.e., school performance test scores, professional networks) and the *context and conditions* are operationally defined as a school-based forum conceptualized as the internal accountability system.

The components of an internal accountability system, to be developed in the conceptual model that follows, incorporate the dimensions of accountability as informed by Leithwood et al.'s (1999) classification schema, Newman et al.'s (1997) structural model of internal accountability system components, and Abelmann and Elmore's (1999) conception of internal accountability.

The Conceptual Model

The study design incorporates concepts from the literature on accountability, assessment and system change. The conceptual model that is explained in this section is used to answer the first research question. The model organizes the analysis of the internal accountability by its five component parts; but it is also used to determine its strength based on the same criteria in the literature. The conceptual framework identifies environmental factors that impact on the accountability policy to practice connections. It

is also used to describe the relationship between external and internal accountability systems. In addition, assessment practice is characterized in order to answer the second research question. The conceptual framework also qualifies aspects of school assessment practice in order to determine school response to the aims of the external accountability system.

To answer the first sub-ordinate part of research question one, an integrated model of internal accountability is outlined. It takes into account both organizational structures and professional interactions that define the context and conditions of a school. The conception of an internal accountability system applied in this study is a composite construction borrowing from the works of Newman et al. (1997) and Abelman and Elmore (1999). The integrated model appreciates parallel aspects from the two models; the former model being a structural conception while the latter a more organic conception. These aspects reflect the underlying epistemology as viewed from different theoretical perspectives and the relationships now presented edify the rationale for the integrated model. The consequences of the former model parallels the feedback inherent in the mechanisms of the latter that leads to development of shared expectations (aims). The constituency (persons) who judges performance in the structural model is related to what teachers' believe about their responsibilities (the conception of responsibility) in the organic model. In other words, who is responsible for making the judgment? This last comparison illuminates the management perspective and highlights the underlying assumptions that contrast with the beliefs about responsibility and authority from a professional (organizational) learning perspective. These theoretical perspectives frame complementary parts for each dimension based on traditional differentiation of authority and power explained in the literature. The integrated model is

a response to an emerging trend in research that advocates for inclusive analytical research models to unpack the complexity of societal organizations and the implementation of change.

The components of the integrated model of internal accountability include: a) the aims, b) the information on school performance, c) the standard of achievement, d) the persons who judge school performance, and e) the mechanisms that guide and monitor school practices. (see Figure 1) For the purposes of this study, the presence of all components is used to flag a strong school-based system, one goal associated with the Ontario accountability agenda. These system components may possess elemental variables that comprise differences in school configurations even when all five components are present and similar in nature.

The first component of internal accountability refers to the *aims* of the internal accountability system. The aims of an internal accountability system are defined in the school mission statement and the goals of the instructional program. The aims include those stated explicitly in school reports and improvement plans but may also include those that are implied in the organizational decisions and routines. The consequences associated with the aims depend upon the level of reporting required by the external accountability system and may involve obligatory explanations or justifications of school programs. Consequences are elaborated in the fifth component because they are inherent in the feedback from accountability mechanisms.

The second component of an internal accountability system is the *information on school performance* gathered as part of the program evaluation inherent in the accounting process. This information may refer to classroom assessment data and grades prepared by teachers for reporting to parents. It may include formalized test scores that

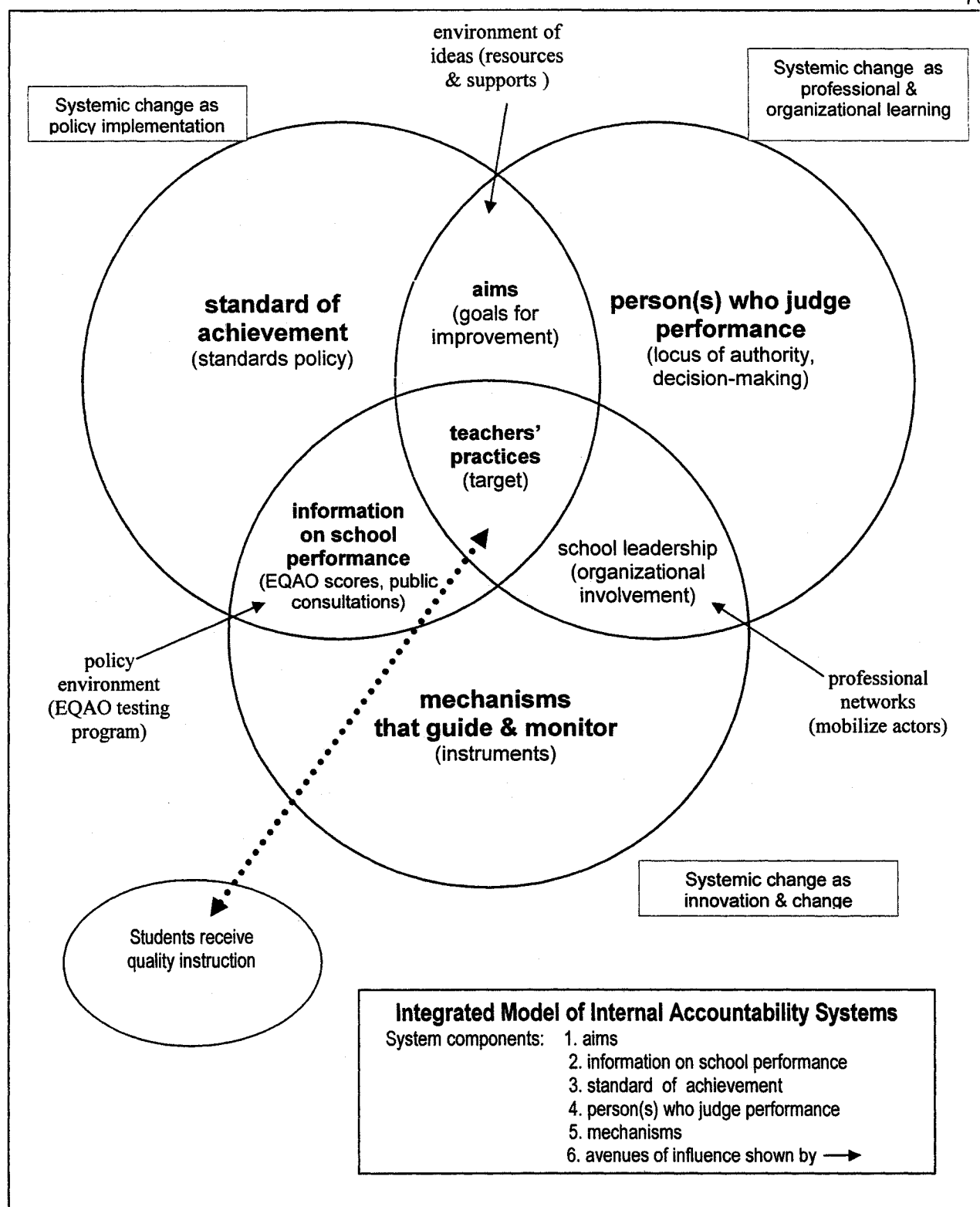


Figure 1. Integrated model of internal accountability systems.

Note. This model is a conception based on the components of internal accountability as interpreted from the literature on internal accountability systems (Abelmann & Elmore, 1999; Newman et al., 1997).

schools collect or gather from external agencies to provide information on achievement, aptitude, or special needs. The information may also include school specific socio-economic or demographic information that can assist with assessing student potential and school performance.

The third component of an internal accountability system refers to the *achievement standard* that is used by teachers and other stakeholders to determine the level of student achievement and the effectiveness of the instructional program. Evaluations of school programs can be based on standardized measures identified by external agencies, but school grading may reflect internal school expectations that may or may not be consistent across classrooms, schools or even districts for that matter. How a school and its staff develop collective expectations and use a common standard of achievement reflects teachers' perspectives about their responsibilities.

The fourth component of an accountability system refers to the *persons* who receive information on school performance, (i.e., the collective student and aggregated school performance information) and *judge* the instructional program based on whether school performance meets school expectations. This person or persons may provide feedback on program, change the instructional program, or require that rewards or sanctions be meted out.

The fifth component of an internal accountability system refers to the policy requirements, organizational procedures, and professional interactions at the school level that operate as *mechanisms to organize, guide and monitor* teachers' practices. Accountability mechanisms are the formal and informal ways that people in schools are accountable to a person or persons of authority. Professional feedback or changes in organizational procedures and practices are linked to such mechanisms. The nature of

the feedback is determined in part by the school aims and collective expectations. Feedback may be viewed as a reward or sanction depending on the intention behind the action and the level of accountability associated with it. Some consequences may be bureaucratic while others may be moral in nature. There are two types of mechanisms and two forms of feedback.

Internal mechanisms may include lesson/unit planning, long-range plans, school reporting procedures, personal growth plans, and verbal monitoring. The design of the mechanism and the nature of the accounting reflect different approaches to accountability and may vary between schools. *External mechanisms* may include large-scale accountability program requirements, student report card protocols, public reports, school improvement plans, target-setting exercises or performance appraisals. School-based mechanisms also possess inherent feedback.

Internal feedback may be organizational (resource distribution, staff duties, reporting procedures) or professional (coaching, professional development, team decision-making) in nature. Internal feedback comes from school reviews, teacher performance appraisals, and collegial dynamics. Associated with feedback are issues of school reputation and professional esteem that relate to meeting the school aims. *External feedback* is the information on school performance (testing results) and recommendations for school improvement planning. Schools receive feedback from district administrators, parents and the public in different ways.

When the school performance data is aggregated and reported publicly in a form that invites comparison with provincial and district performances, this feedback becomes a factor that may help or hinder the effectiveness of the internal accountability program. In addition, the feedback process and how it is handled depends to a certain extent on the

type of leadership within the school. Leadership represents another factor impacting school-based accountability. Such factors are a function of the external policy environment.

The elements that make up the internal accountability systems of the school may create a configuration that is specific to a school, although commonalities are anticipated. This study uses the same criteria as Newman et al. (1997) to qualify a strong internal system. Therefore, the existence of all five components within a school is interpreted as meaning a strong internal accountability system is present. These aspects of the accountability systems are used to answer the first sub-ordinate part of question one.

In order to anticipate the avenues of influence that the external policy environment has on school-based accountability systems, factors that impact on internal accountability systems are the focus of the second sub-ordinate part of question one. For the purposes of this study, these factors are differentiated into two groups (See Figure 2). The first group refers to the mandated requirements associated with policy implementation process. The second group refers to the mobilization of supportive actors who are able to access professional ideas (expertise) and employ leadership practices.

The *policy requirements* of an accountability system structured around a large-scale testing program may include mandates related to test participation, transparent public reporting of school performance scores, and recommendations.

The *mobilization of supportive actors* is another avenue of influence and is linked to the actions and decisions made educational leaders with responsibility for accountability processes. Leadership is one factor that may help or hinder internal

Sphere of Influence	Policy to Practice Avenues of Influence
Policy Requirements	<i>EQAO performance-based testing program:</i> • transparent reporting • program review (assessment and improvement planning) • test participation (results as an external measure) • recommendations
Mobilization of Supportive Actors Professional Networks	<i>Leadership</i> • collegial processes • guidance • allocation of resources <i>Professional interaction</i> • collaboration (networks) • dialogue and discourse <i>Professional Politics</i> • political issues and constraints

Figure 2. *Factors Impacting Internal Accountability Systems.*

Note. This figure is a representation of selected avenues of influence reported in the literature (Chrispeels, 1997; Datnow & Castellano, 2001; Earl & LeMahieu, 1997; Firestone et al., 1998; Fullan, 1991; Knapp, 1997).

accountability processes because it is leaders who press for change and encourage experimentation. The presence of transformational leadership to establish trust in process, facilitate structures that promote professional learning, and access expertise to support internal accountability system processes is an enabling factor. Leadership is a critical element in building collective expectations and establishing mechanisms to promote meaningful change. Instructional quality can be strengthened when action is taken to organize professional interactions, align school professional development with school goals, promote socialization among the staff, and practice shared leadership (Youngs & King, 2002).

On the other hand, change initiatives can be hindered by feelings of alienation when members of the teaching community do not work together in a collective effort. This balkanization can be an individual or a group phenomenon (Fullan, 1991). One leadership function is to promote collegial dialogue and discourse while developing organizational structures to support professional learning. One can expect a subtle balance between contrived collegiality and true collaboration. The former characterized by formal procedures such as joint planning, scheduled meetings, and peer coaching, while the latter involves personal, enduring planning and exchange of ideas and materials related to student development (Fullan, 1991). Leadership is key to owning the processes of internal accountability by adapting policy in meaningful ways and, in doing so, avoid professional dilemmas that Spillane (1998) described as policy adoption and co-optation.

In considering the overall nature of internal accountability systems in an external accountability framework designed to foster its development, the strength of internal accountability is closely tied with its ability to respond to the dual aims of external accountability framework. Compatibility between the aims and between the instruments and strategies of both internal and external systems is explored to determine the nature of the relationship in answer to the third sub-ordinate part of the first research question.

In order to capture the nature of internal accountability systems, this study characterizes school response to the first aim identified in the Ontario restructuring agenda, the accounting for the quality of the instructional program to the public. School response is described in terms of the EQAO recommendations as they pertain to communicating and interpreting assessment results. These recommendations advise schools to: a) consult with teachers, parents and the community to establish action plans

and review them regularly; b) use classroom assessment strategies to determine what students need to learn next; c) identify school and community factors related to student achievement; d) communicate about the purposes of assessment, expectations for students, the quality of student work and home support; and e) discuss achievement and improvement strategies (*Provincial Report on Achievement*, 1997). School response to these recommendations reflects the nature of the internal accountability system and speaks to the first research question.

The second research questions queries the mediating influence of an internal accountability system to influence teachers' assessment practices. To explore this influence, classroom assessment is first described using the characteristics of classroom assessment as described in the previous section and then analyzed for cooperative and collaborative assessment goals, methods and attitudes (see Figure 3). This analysis supports suppositions about whether the second aim of the Ontario's accountability program, enhanced teacher practice, is occurring within schools. The Ontario accountability framework identifies dual aims, school accountability and professional learning. The propositions arising from this review of the literature considered these dual aims and the development of this conceptual framework was guided by these propositions and the research questions.

For the purposes of this case study research, accountability policy is viewed from the position of the standards movement. This study examines the internal accountability systems and their influence on teachers' assessment practices in mathematics. This study was conducted within a provincial environment where systemic change was being implemented and accountability policy reflected a professional practice. The data describe the components of internal accountability systems and the policy to practice

Characteristics	<i>School Level Assessment Dynamics</i>
Goals	Resistance and reluctance to standards-based assessment <u>versus</u> acceptance of assessment policy.
Instruments	Achievement level is assessed by individual <u>versus</u> collective criteria.
	Testing <u>versus</u> active use of portfolio (assessment toolkits) for evaluation purposes.
	Use of commercial tests <u>versus</u> use of standards-based school and curricular performance tasks and tests evaluated with specified criteria (exemplars).
Methods (processes)	Minimal collection of evidence to demonstrate student learning. Assessment is product oriented and usually summative <u>versus</u> ongoing planned collections that use a variety of assessment strategies that focus on process goals as well as end product. Diagnostic, formative and summative assessment conducted.
	Limited use of assessment terminology <u>versus</u> correct use of terms to communicate clearly.
	No collaboration <u>versus</u> discussion of assessment criteria, expectation and the development of assessment tasks and tools (rubrics).
Personnel	Individuals feel isolated from professional colleagues and lack professional dialogue <u>versus</u> school professional dialogue, shared decision-making and consultation with district staff and the school community.
	Individual teacher assessment <u>versus</u> collegial and student involvement.
Value system	Teacher independently assesses end product for grading purposes <u>versus</u> teacher valuing collaborative assessment processes that articulates a learning skill focus and encourages student self-assessment.
	Teacher does not seek support <u>versus</u> school professional staff networking and seeking support from the district staff and local community.
	Classroom assessment focused on summative assessment and content/ procedural knowledge <u>versus</u> a focus on performance tasks, the learning process, variety of strategies and individual learning plans.

Figure 3. Dynamics of School Assessment Practice.

Note. From "Level descriptions and teacher assessment in England: Towards a community of assessment practice," by K. Hall & A. Harding, *Educational Research*, 44 (1), p. 6. Copyright 2002 by Educational Research, <http://tandf.co.uk/journals>. Adapted with permission.

connections that impact on them. The study also explores the relationship between the internal and external systems by describing the compatibility between the aims and the strategies at each level of the system. It also describes school response to the accountability program recommendations.

Lastly, the discussion considers cross-case findings related to the nature of internal accountability systems and their influence on teachers' assessment practices. The study does not focus on issues of professional learning, assessment literacy and district capacity building. These are recognized as significant factors and they were taken into consideration during the selection process of the district and the participating schools. They represent fixed variables in the study design.

Overview of the Chapter

The review of the literature in this chapter developed the study propositions from earlier works in three fields: accountability frameworks, assessment in education and theories of system change. An innovative framework operating in the Province of Ontario was described in order to situate this study of internal accountability systems. The rationale of the research design was introduced and the research questions identified. The conceptual framework was *operationalized* using concepts and models from the literature and the intent of the study was outlined. The method is described next in Chapter 2. This is followed by descriptive case studies of two elementary schools organized as Chapters 3 and 4. Chapter 5 contains a discussion of a cross-case analysis and study findings. Chapter 6 reports on the conclusions, of the limitations, and the implications of this project for education and for research.

CHAPTER 2

METHODS

The purpose of this collective case study was to explore and describe the nature of extant internal accountability systems at the elementary school level and to capture the impact of these systems on teachers' assessment practices in a context of standards reform. The objectives were four fold: a) to analyze school structures, procedures and patterns of behaviour that comprise the internal accountability system; b) to explore the factors that influence internal systems; c) to determine how these entities respond to external accountability agendas that include large-scale performance-based testing programs and recommendations for school improvement; and d) to describe school level response to standards-based assessment policy.

This exploratory study employed a pragmatic design. It used qualitative methods to capture and describe the content, processes and interpersonal dynamics of the internal accountability systems; but it also employed a mixed model approach (Johnson & Christensen, 2004), in the sense that the selection of the participant schools was purposive and the conceptual framework proposed a tentative conceptual model to begin the analysis and interpretation of the data. The qualitative methodology helped to ground the study within a constructivist paradigm because this enabled descriptions and interpretations of the complex world of schools taking teachers' perspectives into account.

This collective case study was interpretive in nature because it aims to explore and describe school organization structures and patterns of behaviour based on interviews, demonstrated in documents, and described in field-notes of observations and communications. The unit of analysis (Merriam, 2002, Yinn, 1994) is the internal

accountability system at each school as perceived by the professional community, specifically the administrators and teachers. Embedded within each school case study is an examination of teachers' assessment practices. Analysis of the organizational structures, processes and patterns of behaviour framed the components and characteristics of the internal accountability systems and shaped the descriptions of classroom assessment practices and school response to external accountability policy aims and mechanisms.

The case studies were instrumental in design (Stake, 2000) because the objective was to characterize the complex internal accountability systems using and building upon the concepts, models and theories in the literature. The study was also collective in design (Stake, 2000) because two schools were examined in order to determine if the components of internal accountability systems were common in both sites and to note similarities and differences in the school configurations. The replication logic in the study design enabled some exploration of possible school level response and helped to frame tentative suppositions about the nature of internal accountability systems found within an external accountability system that valued professionalism.

The data for this study were collected as the schools were immersed in implementing standards-based assessment within an external environment of province-wide performance-based testing. The study focused on two themes, firstly, the nature of the school internal accountability system in such an environment, and secondly, the influence of the internal accountability system on teachers' assessment practices. The study also sought to move beyond the initial reactions of teachers that reflected their anxiety associated with the unknowns of the accountability movement and their frustration with the practical concerns and dilemmas present in contexts of systemic

change. These significant drawbacks are well documented in the literature and were expected within the findings of the case studies as well. The intent was to explore the role of the internal accountability system to enact policy to practice and describe the connections present.

This chapter explains the selection of the study participants, describes the nested accountability policy context, identifies the data collection and analysis methods, and speaks to the trustworthiness of the data.

Participants

The participant schools were situated in a nested policy context. The provincial and district policy contexts are described in order to understand the external environment for each case. General characteristics of each school are provided next in terms of size, community factors and program options. Pseudonyms are used for the district and schools, and all persons participating in this study are identified by their area of responsibility (i.e., Grades X teacher).

The selection of participants for this study of school internal accountability systems involved a number of important criteria. A provincial context involved in systemic reform using large-scale testing for accountability purposes was first considered. The nested policy context of the school within a district and provincial environment meant that district selection was an important second step. Based on research, strong internal accountability is associated with high capacity but the reverse relationship is not a given. In order to examine internal systems, it was advisable to select a district with support staff and an active professional development program, two indicators supporting high capacity. A third consideration was to invite schools

interested in examining their mathematics assessment practices and seeking to explore school dynamics related to assessment practice so that participants were prepared to commit the time and effort required by the study. To address internal accountability systems, the decision was made to study middle schools (K to 8) because province-wide testing had been established at the grade 6 level for three years. Teachers and principals were familiar with the performance tests and the trends in the assessment data. In middle schools, School Improvement Plans [SIP] were required and EQAO recommendations were available. Specific requirements related to the teacher participants were a function of the grade level taught as personnel characteristics or teaching experience were not deemed relevant. The selection process and the context for each level of participants, including teachers, are now described.

Province

The Province of Ontario began changes in its educational system in the 1990's, as it moved towards standards-based policies. The Ministry of Education and Training called for a new approach to curriculum and instruction, standards for student achievement, and the establishment of an accountability system centred on a province-wide student assessment program.

The new Language Arts and Mathematics curriculum initiated in 1997 was more specific than previous curricula regarding both the knowledge and skills that students were expected to acquire in each grade. Within the language and mathematics curriculum for each grade, the subject material was organized across key domains, with specific outcomes for knowledge, skills and applications listed. Exemplar documents were prepared that showed the characteristics of student work at each level of

achievement for each grade. The standards of achievement for each grade level possessed four categories. Level 3 was considered the provincial standard and student performance was compared to this standard as qualified by the criteria in the achievement charts found in the curriculum documents. These resources were designed to promote consistency in the instructional programs and the assessment of student achievement from grade to grade across the province.

The Education Quality and Accountability Office (EQAO) was established in 1995 by the Government of Ontario to ensure greater accountability and contribute to the enhancement of the quality of education in publicly funded schools through its large scale testing program. It audits school and district performance through its assessments, questionnaires and improvement planning reviews. It is responsible for objective and reliable reporting of quality indicators and testing results; and for making recommendations for system improvement.

Provincial assessments are conducted in reading, writing and mathematics at grades 3 and 6. A third mathematics assessment test is administered in grade 9, while a literacy assessment is conducted in grade 10. These assessments are mandated for all students. These tests involve multiple-choice questions, short answer questions and extended performance tasks. Districts receive district, school and individual student results that are forwarded to the schools. School results and district results are sent to the schools for staff and Parent Council review to support discussions about the development of school improvement plans. At the district level, the school board is responsible for having school and district aggregated results published for public review.

The Education Quality Indicators Program (EQUIP) is an informational tool developed and overseen by the EQAO. It contributes information on non-academic

factors that influence student achievement. These data are also used to inform district and school improvement planning. At the district level, staff work with teachers, administrators and parents to develop local school improvement plans (SIP). The school districts themselves also develop a district improvement plan (DIP) based on a template provided by the EQAO. The district level plans are submitted annually for monitoring and feedback.

The EQAO mathematics tests, in particular, are developed via a collaborative process among educator groups and assessment specialists within the province (EQAO officer, November 2002). The EQAO staff develops a blueprint that reviews the five mathematics strands, the expectations and the categories in the achievement charts and identifies the curriculum targets that the annual assessment tasks and questions will address in each test. In Ontario, the five mathematics strands addressed are: a) number sense and numeration, b) measurement, c) geometry and spatial sense, d) algebra and patterning, and e) data management and probability. Educators collaborate on the development of the test questions after they have received training on test item development. The items are first used with a small sample of students and then revised. Field-testing of the questions is done with a large sample of students designed to be representative of differing locations, genders, and language. The short answer and performance task components of the tests are changed each year, while the multiple choice test remains the same.

Performance standards for each test are derived through a standard setting exercise. This process examines a range of responses from the field-testing. A scoring guide is constructed from the analytic traits used to develop the test items, then trait marking procedures are used to score test items. Representative samples are taken from

the students' responses for each of the four achievement levels based on the criteria on the appropriate grade level achievement chart. Level three is considered the provincial standard and may be referred to as the benchmark standard. These anchor documents are used in the training sessions for markers. The tests are marked by teaching professionals during the summer vacation period. The process is monitored and moderated. Assigned scores are cross-referenced and rechecked for validity before the final evaluation is determined.

The EQAO administers other national and international tests to randomly selected samples of students from across the province. The School Achievement Indicators Program (SAIP) is an instrument of the Council of Ministers of Education, Canada; while the Third International Mathematics and Science Study (TIMSS) was administered in 1995 and 1999 under the auspices of the International Association for the Evaluation of Educational Achievement (IEA). The main purpose of these assessments is to permit national and international comparisons of achievement, teaching methods and curriculum among participating countries and provinces.

District

The Central School Board (a pseudonym) was selected for this study because it possessed an extensive district support staff and an established organizational structure able to develop and support staff development programs, system reviews, and improvement planning. The district administration was recognized for its leadership in the province (District Officer, August 2002) on school improvement planning. In addition, it was a sufficiently large district that it possessed several elementary schools. It also had a formalized process for applications to conduct research.

These district characteristics helped to narrow the variability for three of the four essential elements within Knapp's policy to practice framework (policy environment, organizational capacity, avenues of influence to inform school and district capacity building). These characteristics were common to both schools in this study. The significant variables in the study design were functions of the component parts of the school internal accountability system. For the purposes of this study, these were the context and conditions of the school organization and its professional community.

The Central School District represented a rich ethno-culturally diverse community. The mission statement reflected this in the phrase, "challenges all students to achieve personal excellence in learning and responsible citizenship within a safe, equitable, diverse and caring environment."(*District Strategic Plan 2000-2003*)

In September 2000, the district Trustees adopted the strategic plan. The renewal plan was developed from a collaborative process involving students, parents, community stakeholders, trustees, teachers and district staff. It identified nine areas for improvement for excellence in learning. These included: a) expectations for learning and accountability; b) staff training development and support; c) individualized learning opportunities; d) safe, caring schools; e) school as hub of the community; f) volunteerism; g) boards structures and processes; h) corporate management; and i) fiscal planning.

The goals of the district strategic plan were operationalized as twenty-two action plans that detailed a concrete set of improvement goals for the system and all its schools. The performance measures for each plan were monitored and tracked in order to provide evidence of improvement during the 2001-2002 school year. The results were intended as indicators for continuous school improvement planning. Two highlights arising from

the progress report were: first, the commencement of achievement targets for all secondary school graduates and Grade 1 students; and second, the need to inventory technological resources for improving the effectiveness of the teaching process.

To address issues of program evaluation, district staff developed four school manuals. These included guidelines for the school review audit, a manual on school improvement planning, a practical guide on assessment and grading, and a reference document on the use of statistical information including EQAO test results.

Superintendents initiated a school review process where regular visits to each school were designed to assist principals identify strengths and weaknesses in each of a number of areas in order to assess what aspects of school teaching, learning and operations need additional focus and support. This process guided principals as they formulated strategic management plans and initiated school improvement planning. SIPs were linked to system-wide goals by district literacy and numeracy initiatives.

District personnel were responsible for liaising with the EQAO and for working with program support staff. They also liaised with Board trustees to develop the system improvement plans (DIP). District staff also worked with school administrators and teachers to develop SIP, organize information on school and community characteristics, and analyze EQAO data. At the time of this study, the District was implementing annual SIP workshops. No formal process was in place to provide feedback on SIP implementation.

Schools

The following criteria are purposive and were used to select participating schools:

1. the schools must be situated in a nested policy context of standards reform with a large-scale testing program;
2. the schools and school district must possess strong organizational capacity and expertise;
3. the school must have intermediate and junior level divisions; and
4. the schools must volunteer to participate.

To meet the second criteria, strong organizational capacity was defined as the ability to provide quality instruction. In educational studies (Corcoran & Goertz, 1995), capacity building is associated with strategies that increase potential (*qualified staff, professional development, access to knowledge*); focus on instruction (*standards, goal-setting, grading policy*); and mobilization of teachers (*collaboration, shared decision-making*). To identify the selection of schools for the study, the researcher provided the above criteria to the district officer responsible for school access. Five district schools, identified as suitable by the quality assurance officer, were contacted and study information distributed. Due to reasons specific to each site, two schools expressed interest and later committed to participating in the project. Reasons for declining participation from the remaining candidates included a new administrator in the school, math assessment practice not being one of the specific concerns currently addressed by staff, and reluctance to have a researcher in the school during a politically tense period related to funding and school closure issues. Ideally, the study called for three schools for replication reasons, but willing participation by two qualified schools was deemed adequate to meet the objectives of this exploratory study into the nature of internal accountability systems.

Each school that chose to participate possessed kindergarten to Grade 8 classes. While the two schools shared some characteristics, they differed with respect to location, socio-cultural makeup, demographics and some instructional programs.

Southview Elementary School

Southview Elementary School is located in an established suburban residential community. The school is situated in a largely bilingual community and offers four programs options: regular English with Core French; Early French Immersion; Late French Immersion (grades 7/8); and an Instrumental Music program (grades 7/8). The school has a Special Education Learning Centre (SELC) with one resource teacher who provides remedial support and educational assessments to its population of approximately 400 students. Some English as a Second Language (ESL) support is available through a family reception centre located within the school.

Southview School serves a middle class socio-economic population and offers a variety of clubs and school enrichment activities to its students. It benefits from an active parent community that fund raises, organizes and operates school projects (i.e., school yard project), and supplies volunteers in the school.

School results from both the Grade 3 and Grade 6 EQAO testing show that the percentage of students performing at levels 3 and 4 was above both the Board and Provincial averages. In mathematics at the Grade 3 level and in reading at both the Grade 3 and 6 levels, the school scores were consistently higher than district and provincial results based on aggregated averages generated by the district staff and reported in the school profile letter sent home to parents.

Northside Elementary School

Northside Elementary School is an urban school located in a culturally diverse community. Many families are recent immigrants and many students speak little or no English in the home. Approximately 25% of the 650 students are English as Second Language (ESL) learners. The school offers a regular English Program with core French and a congregated enrichment program. In addition, the regular program is supported by a special education team composed of one special education teacher (SERT) and three learning centre teachers (SELC). Two full time ESL teachers support the ESL students through a withdrawal program. It also houses one system special education behaviour class.

Northside School was identified as a Signal School by the District, which means that it has students who are considered at risk for school failure. The district used an index to identify such schools, based on factors related to income, housing, family unit, absenteeism, transience, language, culture, and readiness to learn.

Students at Northside School came from a broad geographic area because it offered enrichment and behaviour support classes and these students were bussed in. The local area in which the school is situated possesses properties that range in value from low to high income, although the school population was weighted to the former. Thus the population was diverse socio-economically as well.

School results for the Grade 6 EQAO testing indicate that the percentage of students achieving at Levels 3 and 4 in reading, writing and mathematics was beneath both district and provincial aggregated averages. The same trend was evident in the results for the Grade 3 testing. A comparison of the Grade 6 aggregated average with the Grade 3, revealed improved numbers of students meeting the level 3 and 4 achievement

levels. This suggests that student learning, as measured by the EQAO tests, showed improvement between the testing years and may well reflect the benefits of the instructional program and students experience within the school.

Teachers

The principal at each Kindergarten to Grade 8 School identified teachers willing to participate in the semi-structured interviews. It was requested that the teachers be representative of three grade levels (5, 6 and 7) and have differing responsibilities within the school. These grade levels were requested because they represented junior and intermediate classes that may be most interested and affected by the Grade 6 EQAO testing. One intermediate (Grade 7) and two junior classroom teachers (Grade 4 or 5 and 6) were interviewed at each site. One other teacher, identified as a school leader with responsibilities in special education, was also interviewed at each school. The principal assisted by approaching teachers who were participants on the SIP committee, so that some participants were knowledgeable about the SIP processes and procedures. At each school 4 teachers and one principal were interviewed.

At both schools, the principals were possibly gatekeepers in terms of which teachers were first approached to participate. The requirement that specific grade levels be represented helped to open participation to a broader group. Since both principals advised that at least one teacher participant possessed strong reservations about the EQAO testing and the provincial standard of achievement, this researcher considers that a range of perspectives and attitudes are represented in the case study descriptions. It is fair to also note that schools volunteered to participate and this may indicate some narrowing in professional perspectives.

Data Collection

Data for this *collective case study* was gathered using a *backward mapping* approach. This procedure (Elmore, 1990; Yinn, 1994) was chosen to facilitate understanding of school assessment practices and accountability practices and forums from the teachers' perspective first, then bringing in school leader explanations. In examining the structures, processes and patterns of behaviour as explained by classroom teachers, an unfolding picture of the components of internal accountability systems was studied and the different areas of responsibility clarified. This procedure was facilitated by managing the data collection using four phases starting with the classroom teacher, then school administration, followed by district and then provincial data. As anticipated, this enabled a focus on policy to practice issues beginning where the policy is enacted and tracing back to where it is developed. The focus was to explore the school infrastructure and patterns of professional interaction responsible for the policy implementation. It also permitted participants' awareness of the policies and system forums to be identified. This section is organized to outline the instruments used and the procedures followed in the data collection.

Instruments

The instruments used to collect data for the study includes semi-structured interviews, document collection, and field notes. Details regarding each instrument are given below.

Semi-Structured Interviews

Protocols were used to guide the direction of the semi-structured interviews.

These protocols were organized into two parts, the first part possessed questions related to assessment practices while the second part focussed on eliciting information about accountability system components and professional interactions within the school community environment. The protocols used in the study are found in Appendix F.

Each protocol (Yinn, 1994) contains a set of substantive questions organized according to the dimensions presented in the theoretical framework. To guide the focus of the semi-structured interviews, each substantive question clarifies the concepts, structures, procedures, and issues to be explored and described (i.e., assessment data, teacher practices, decision-making responsibilities) during the interview. As study participants are unlikely to recognize the innovative forum that constitutes internal accountability, the questions must be reframed by the researcher as related questions narrowing the inquiry to routine procedures, organizational structures and professional interactions. For example, the question “What are the specific components of the school’s internal accountability system?” May be initially rephrased as follows:

1. Do you track student assessment data to get an idea of overall school achievement?
2. Are school (students) EQAO performance results compared to previous school scores? Compared to aggregate district or provincial scores? What actions are taken as a result?
3. Do staff members discuss the expected level of achievement? Who determines if the current school scores are satisfactory? How is this done? Who discusses what changes are needed? Who would be responsible for carrying out action plans?

4. Can you explain what type of follow-up occurs when school performance is reviewed? Give an example of something that has happened recently in this regard.

These questions lead to teacher perspectives and explanations of assessment practices, school responsibilities and program evaluation; and then to descriptions of school routines, requirements, and improvement planning.

Additional questions employed the language of the interviewees and led to further clarification and descriptive data. Information about assessment practices, use of assessment information, ownership for school improvement were of particular interest, however, the impact of environmental factors on the internal system was also examined. Emerging lines of inquiry teased out details of *how* schools (teachers') made an account of their teaching and assessment practices, *what* they saw as their responsibilities, and to *whom* they reported. Furthermore, issues arising from the responses were captured, some followed up immediately while others were pursued at a later point. Through a conversational dynamic, the components of the internal accountability system and its relationship to the external system were elaborated. The questioning usually concluded when common themes were noted.

The interviews were audio taped for two reasons. One was to enable the interview to unfold in a conversational manner that enables subsequent questions to follow logically from preceding responses. Furthermore, it permitted the researcher to seek clarification and detail on processes, procedures, relationships and patterns of behaviour with the school organization without need for extensive note taking. Audio taping ensured accuracy of what was said in word and tone and assisted with the transcription process. The audiotapes, once transcribed, were stored as text files for use

in a qualitative software program (N-VIVO). These were also printed as hard copy and used for initial analysis of the data.

Document Analysis

Documents were gathered to provide information on educational policy, school review processes, school improvement planning requirements, accountability reporting and classroom assessment practices. The collection was organized into three categories: a) school site documents, b) district publications, policy and reports, and c) provincial accountability reports, policy guidelines and recommendations. The school site documents were separated for each location. School documentation included samples of teacher assessment instruments and tools, school profiles, SIPs, and other relevant information provided by school staff.

Field Notes

Throughout the study, field notes of descriptive and analytic observations were made during interviews, discussions, telephone calls, school visits, early stages of data analysis and email communication. These were an initial level of data analysis that helped to guide the focus of upcoming interviews so that early information could be clarified and interesting areas for further inquiry could be followed.

Procedures

The data collection was managed using four phases that began with data collection at the classroom level, followed by data collection at the school level. After the school data collection was completed, data were collected from participants at the district, then provincial levels. Each phase is further described in the next section.

Phase 1 involved interviewing classroom teachers in grades 5, 6 and 7. This phase began with a brief meeting with all school level participants to discuss the procedures, answer questions and sign letters of consent (see Appendices A, B, C, D and D for letters of consent). This meeting was followed by two semi-structured interviews approximately two months apart for the classroom teachers. The first interview focused on mathematics assessment practices. The second one solicited perspectives and descriptions of the school improvement planning process, school level accountability mechanisms and school professional community interactions, in other words, the internal accountability system. Helps and hindrances in the implementation of standards curriculum and assessment were discussed. These data related to research Question 1 (a and b) and Question 2.

Phase 2 involved interviews with the school principals and lead teachers. In both cases, the special education teachers also participated. These interviews focussed on the *unit of analysis* for the study, the school internal accountability system, and its professional community. The principal was the last staff member interviewed. This second phase helped to detail the components of the internal accountability system, further explore the factors that impacted on it, and link the interplay between internal and external mechanisms from a school wide perspective. This phase contributed data for Questions 1 (a, b) and Question 2. It also provided specific information related to research Question 1(c).

Phase 3 and *Phase 4* involved consecutive interviews with a district officer and an officer from the provincial EQAO to address research Question 1 (c).

In addition to the interviews, documents were collected at each phase of the study. Samples of teacher assessment tools were collected to support teacher

descriptions of assessment practices during Phase 1. School letters, profiles, guidelines, SIPs and information related to the school context were collected during Phase 2. District profiles, reports, DIPs, policy manuals, and relevant resources were collected during Phase 3, and in Phase 4, provincial policy documents and publications were gathered from websites. The study was conducted over a time period of approximately one year beginning with the submission and review of the application to conduct research and continuing until the last interview was conducted. The active part of the data collection in the schools took place over a six-month period.

Seven interviews were conducted at each school, ranging in length from 50 minutes to approximately 2 hours depending on the interviewee's availability and willingness to provide in-depth descriptions. Most interviews were 75 minutes in length. A total of 16 interviews were audio-taped and transcribed. First, teachers were interviewed about assessment practices beginning in January and early February. A second interview was arranged after a reporting period and parent-teacher interviews were concluded. In phase two, the school resource teachers were interviewed for the second purpose, followed by the school principal.

Contact with the school participants continued until the condensed summaries of the interviews were member-checked (Stake, 1995) and additional voluntary comments incorporated. This process overlapped somewhat with the interview of the district administrator, and later this overlap occurred with the EQAO representative. This was an anticipated event and did not detract from the study design. The main qualitative analysis began with the compilation and organization of the data from Phase 1 and 2, and continued as Phase 3 and 4 were completed.

Data Analysis

Policy to practice connections are complex multi-faceted relationships and the analysis of these relationships involves multiple dimensions. The analytical framework for the study helped to identify the different analyses needed. To examine the components of internal accountability, a microanalysis of school-level processes, procedures and professional interactions was needed. To clarify relationships between the internal and external aims and mechanisms for example, a macro-analysis of the dimensions of accountability and the role of environmental factors was performed. The analysis for each research question involved three actions: data reduction, data display and interpretations (Hubermann & Miles, 1994).

A qualitative analysis of the text was completed using computer software. This preliminary stage was later cross-referenced with information gathered by the review of field notes and the documentation (i.e., teacher assessment tools, school and district plans, publications, policy manuals and resources available to educators) to support further analysis and data interpretations. Each facet of the process is further described in the following paragraphs.

The main data reduction of the transcribe text used a qualitative software program, N-VIVO, to organize and compile instances according to start codes arising from the conceptual framework. Each school site was coded separately as each case was a separate unit of analysis. Next, the text from district and provincial interviews was coded. This process commenced when the transcribed text was member-checked and converted to *rich text files* (rtf.)

The text was read and passages that referred to the first research question, “What are the characteristics of school internal accountability systems?”, were *visually coded*

into three broad groupings based on the subordinate parts of the first question. Text that spoke to internal accountability system components was coded green, text that spoke to different avenues of influence was coded blue, and text that related to external accountability policy or related policy mechanisms was coded red. Text that addressed the second research question, “What is the influence of internal accountability, if any, on classroom assessment practice in a context of external accountability?”, was coded orange. The four categories were assigned *start codes* (IntAcct, Factors, AcctFr, MI) and were identified as *tree nodes* within the software data management system. Within each coloured section of text, second order start codes were used to organize the data analysis by specific concepts and relationships anticipated in the conceptual model. As the passages were reread, text specific to the components of the internal accountability system was identified and allocated the appropriate letter code.

The second step involved analysis for specific components and classroom assessment characteristics. This *micro-analysis* identified textual references that related to the components of the internal accountability as defined in the conceptual framework. Characteristics of teachers’ assessment practices were also flagged and coded. These sections of text were then re-sorted to *compile instances* displaying common concepts, relationships and themes. A list of start codes and final codes is found in Appendix G.

Rereading the compiled text for each of the 4 original categories (tree nodes) led to *coding on* that developed *sibling nodes* and *child nodes*. This process meant that passages assigned to internal accountability (IntAcct) by the initial visual coding, were reviewed for specific concepts and relationships outlined in the conceptual model such as information on school performance (sc), standards used for assessing school performance (st), persons involved in judging school performance (ag), and mechanisms

that guided/monitored school practices (cq). As each instance was highlighted and recoded, the text was recompiled into new groups by the software program and labelled as to its relationship with the superordinate category. This process was recursive as reflective analysis and further reading informed the analysis. As sub-ordinate categories emerged from the recompiling, new child codes were developed. Where overlap occurred, the text was dually coded. For example, some mechanisms were further classified as reorganization (reorg) and others as improvement plan (SIP).

Macro-analysis involved clustering and reorganizing coded text as pattern matching and overlaps of coded text revealed relationships and themes. This third round of coding on teased out instances that represented triggers for change (tri) and compatibility between strategies and policies (al). New nodes were created as needed. For example, internal accountability mechanisms (cq) were found to include school-based mechanisms like reassignment of teaching duties (reorg) and external mechanisms like public reports public reports (EQAQ requirement). In some instances, coding was revisited and coded text was deleted, linked, or shifted to more appropriate node locations.

The compiled text was revisited and informed by further reading of the literature and by review of the field notes and document collection. Patterns in professional interactions were matched and contrasted to theoretical constructs and relationships between policies and practice framed.

The *document review* began with a content review of the assessment tools; then continued with an examination of school and district improvement plans, organizational profiles and newsletters, policy statements and guidelines, and other information sources, including web pages. Provincial documents were included in this process as

well. The documents were read and data categorized using concepts and themes from the literature.

Data displays organized the analysis of the documents by key concepts, models and themes for each of the guiding questions. For example, district policy statements and information profiles were read; and the aims, instruments and strategies associated with accountability were identified. These were then organized in chart formats that later became the figures used to display the findings. A similar process was used for provincial and school documentation.

The dimensions of the internal accountability systems were coded into nodes that represented the five components of the integrated model. Once instances for each dimension were compiled, these informed the written descriptions and provided data that was displayed in reduced form in the figures. Data from the documents and the compiled text were used to draft the case study descriptions.

These review processes both informed the data reduction and helped construct the descriptions of the accountability processes and instruments. To these analyses, the information gathered as field notes was added and *cross-referenced* with the findings of the qualitative analyses of the text and documents. These notes together with the coded text and document analysis informed the descriptions of the policy to practice connections and professional interactions that comprised the internal accountability systems.

The analysis of the interview transcripts, the document review and the cross-referencing of the field notes were used to organize the data that informed the study findings. Together these processes provided detailed data that formed the basis of the descriptive results organized as *case study descriptions* in Chapters 3 and 4.

The cross-case analysis of the findings for the two case studies was described and interpreted using the same organization of the guiding research questions. Some observations of the cross-case findings were quantized (Johnson & Christensen, 2004) to capture the presence of specific characteristics within each case study and enable suppositions related to the nature of internal accountability systems and their influence to support collective assessment practices.

Trustworthiness

Within the study design, several strategies were used to gather extensive data upon which to base interpretations. In order to provide triangulation in the collection and analysis of the data, the following were included: multiple methods, relying on multiple informants at each school, and actively seeking out divergent data such as differing viewpoints and descriptions of processes. In addition, member checking was done to enable participants to clarify that the data clearly reflected their intended explanations and perspectives. These four approaches are described below.

The study design employed *multiple data instruments* including interviews, field notes, and document review. Combining and contrasting data from each source provided a balance and check to the analysis of the information and synthesis of the interpretations and findings of the study.

To develop comprehensive descriptions of school internal accountability, data was gathered from five participants at each school. The use of multiple informants, each with different roles and responsibilities within the schools, helps ensure that the accountability structures and programs were described from different perspectives. For these reasons, three classroom teachers, a teacher leader, and a principal were

interviewed at each school. With the same rationale, district (quality assurance officer) and provincial (EQAO quality assurance auditor) staff were also interviewed. It is noted that principals were asked to identify and invite a parent leader to participate in the study but none were identified. One principal explained that the parent council was supportive but not interested in that level of commitment to school operations and the other implied a level of discomfort with the idea at the present time. The study of mathematics assessment and EQAO programs were framed as a professional responsibility.

The third approach used to enhance trustworthiness was *divergent data*. As a practicing classroom teacher, this researcher was familiar with provincial and district level procedures associated with EQAO testing and school improvement planning. Participation in the EQAO summer marking session for mathematics (Grade 6 performance assessments) assisted in understanding the depth and breadth of policies and procedures related to the standards based accountability agenda. Participation in the district SIP workshops also enabled an appreciation for processes in the nested district policy context. This background provided knowledge of many aspects of the research context and when initial interpretations were discussed and unpacked with my advisor, familiarity was balanced with a sense of objectivity. Further data was gathered from both district and provincial level representatives. These three divergent perspectives assisted in ensuring trustworthiness.

As the study took place over the course of a six-month period, the schools were visited several times to meet with each participant at his or her convenience. The backward mapping process permitted the data collection to begin with the specific (the classroom teacher and their understanding of standards-based assessment) and build to a broader picture of the accountability phenomena one step at a time from the individual,

to the school community, to the school leadership, and, eventually, to the district and provincial policy makers. Each successive level was informed by the previous.

In preliminary discussions with each principal, requesting a teacher participant from the grade 6 level and the two grade levels closest to this assessment year meant that participant selection at the school level was purposive and not entirely at the discretion of the principal. This proved helpful in bringing contrasting views to the study as following initial discussions with both principals, they revealed that not all teacher participants fully supported the EQAO testing initiatives and it was expected that contrasting views should be anticipated.

Finally, additional support for the data analysis and findings was made by revisiting the current literature; comparing and contrasting interpretations of the study data with other findings of investigations on similar topics.

Member checking was used to confirm that the intended meaning of interview conversations were accurate and as intended by the participants. Since the transcripts were condensed to clarify meaning and tighten up the sentence structure often natural to conversation, the interviewees reviewed their specific transcript for accuracy and for content. This checking, which is the fourth approach to enhance trustworthiness, provided all study participants an opportunity to clarify specific points and comment on the overall intended responses. Due to concerns about anonymity expressed by individuals within each school, drafts of the chapters were not distributed for peer review.

This chapter presented the methods used in this interpretive inquiry into the nature of internal accountability systems and classroom assessments practices. The research design was explained in terms of approach, participant selection, data collection

and analysis methods. It concluded with an explanation of the strategies used to ensure trustworthiness.

The following two chapters, Chapters 3 and 4, present the analysis of the case study data and summaries of the findings. Chapter 5 discusses the cross-case findings. The final chapter, Chapter 6 highlights the tentative conclusions of this collective case study, speaks to its limitations and then addresses implications for practice and for further research.

CHAPTER 3

THE CASE OF SOUTHVIEV SCHOOL

This case study of Southview School's internal accountability system explores the strength of the internal accountability system and its effectiveness to respond to the dual aims of the EQAO accountability agenda. To this end, this chapter is organized into two main sections framed by the guiding research questions. The first section describes the internal accountability system and determines its strength and compatibility with the external system. The second section characterizes classroom assessment and explores the influence of the internal accountability system to promote standards-based assessment practice.

The principal at Southview School was Mrs. Long. She was a first year principal but had been the vice-principal at the school the year before. She facilitated the participation of four teachers in this study. The teacher participants at Southview School included the Grade 7 teacher, the Grade 5/6 teacher, the Grade 4/5 teacher and the special education Resource teacher.

The Internal Accountability System at Southview School

This section provides answers to the three sub-ordinate parts of the study's first research question. First, the components of the internal accountability system are characterized and the strength of the internal system qualified. Next, the factors that impact on the internal system are described. Following this, the relationship between the internal and external accountability systems is examined in terms of the compatibility between the aims, instruments and strategies.

Strength of Southview School's Internal Accountability System

This section reports results corresponding to question 1a of the study. The findings are organized by the five components of the integrated model of an internal accountability system and are encapsulated in two charts (see Figure H6).

Aims

The first component of an internal accountability system is the aims of the school. This component is represented in the school mission statement, school and community expectations, and leadership initiatives at Southview School. These aims reflected what the teachers and the principal were held accountable for and determined school priorities and expectations.

At Southview School, these aims were woven within the principal's comments and formulated in the annual school report. In discussing her strategic plan, Mrs. Long shared her vision of an effective, child-oriented school that aimed "to provide each learner an education of the highest quality in a positive environment" (School Report, 2001-2002). This aim was achieved by staff commitment to providing quality educational experiences and opportunities to extend interests, talents and leadership abilities. It was supported by parent involvement in consultations during the school improvement planning process, efforts to improve the schoolyard, and collaboration with staff to raise funds for enrichment activities and promote a positive school climate. These aims emerged during the school review audit and were further defined in the goals of the school improvement plan (SIP). These aims and subordinate goals helped the principal frame priorities and program outcomes.

Mrs. Long identified her two responsibilities as working with Board initiatives and developing the SIP. The first one was a function of her administrative role and the

second a function of her leadership role. This professional obligation was her motivation but also she expected the same commitment from the teachers on her staff. She explained that teachers have to be accountable for “constantly reassessing what you are doing and thinking about it” as modeled by one exceptional teacher who despite 33 years experience reassessed what she did everyday and “that’s what makes her excellent” (Principal, Ju02).

As principal, Mrs. Long sought to redefine her role on staff as an instructional leader and mentor. She intended teachers to differentiate their responsibilities too and take ownership for the quality of the environment by thinking about more than presenting curriculum alone. This expanded view of accountability was sparked during the school review audit the previous year. From the review’s consultation process, Mrs. Long felt she was authorized to proceed with implementing changes to address school needs that included staff renewal, organizational restructuring and professional learning.

Information on School Performance

The second component of an internal accountability system is the presence of information on students’ achievement and school performance. The data collected at Southview School was extensive and included student grades, aptitude test results, achievement test scores, and socio-economic indices (Principal, Ju02). The information also included the standardized testing results from the provincial EQAO performance tests. The principal and her administrative team gathered this information together and used it to inform their assessment and evaluation of the school instructional program (Resource teacher, A02).

Central District collated many of the standardized testing results and provided socio-economic data on the student population and school community for Southview

School. The district database held school results for the Canadian Cognitive Abilities Test (CCAT) and the Early Development Inventory (EDI), testing conducted to screen for special needs students in the primary grades. The district distributed data for the math diagnostic test administered to all students in Grade 7. They also provided statistical analysis of EQAO performance results using school, district and provincial aggregated scores.

The EQAO test scores were external measures of students' performance, and Mrs. Long and her administrative team reviewed this information carefully. They recognized that some children performed at their full potential in such situations but others did not. Some variability from classroom performance was to be expected "keeping in mind that it is only one week of the students' school year... you hope that most of the time it's a fair, accurate representation, but if it is not, the teachers will let us know" (Resource teacher, A02). Although EQAO testing results indicated Southview School performance ranked above district and provincial averages, Mrs. Long used other indicators to determine what her expectations for school performance and student achievement should be.

Mrs. Long met with the principal and teachers from the local high school and spoke with former parents. These discussions provided feedback on former students' achievement on the Grade 9 EQAO mathematics test and raised questions about student preparedness.

We found that we haven't been very successful in the math area. So it would appear that our intermediate kids are not ready for Grade 9 math, although our EQAO testing scores are very high at the grade 6 and grade 3 levels. So there is something happening there, so that's what we are looking at changing at the intermediate level. (Principal, Ju02)

When Mrs. Long shared the EQAO results with her Parent Council, this provided a venue for qualitative feedback on school performance. These discussions triggered a parallel review of the Grade 7 diagnostic test results that apparently flagged “last year and again [in] this year’s tests that there was one certain strand that wasn’t being met at all...and that was the beginning of the realization that we needed to change some of the textbooks that we were using and some of the ways we were teaching math” (Principal, Ju02).

Mrs. Long found that this collection of information enabled a good picture of school performance to emerge from a standardized testing perspective, but she was also interested in how her staff assessed and graded students. She had teacher report card grades for all math strands electronically tabulated to compare “how many students received a level four, how many students got a two, a three, a one, so that we could see where the kids were at” (Principal, J02). Using computer technology, Mrs. Long was able to cross-reference grades with EQAO scores. She explained that this information was collected primarily for her own use, but that she would be willing to share it with interested teachers. She felt that not all staff would feel comfortable with this comparison. However, she did discuss individual student scores with teachers to determine the reasons when outliers appeared or when groups of students performed well (Grade 6 teacher, A02). Mrs. Long used all these data to evaluate program effectiveness and to plan for school improvement and professional development.

Mrs. Long and her administrative team formed the SIP committee that was responsible for analyzing all the available performance data and conducting the program review process. This committee facilitated the discussions of EQAO results, student achievement data and other indicators of performance. Trends in the data were noted and

teachers were asked to think about program gaps and student instructional needs when deciding on strategies for the SIP. So part of that process was the EQAO scores (Resource teacher, A02) because it provided an external measure that enabled comparisons and trends in school performance to be studied.

Standard of Achievement

The third component of an internal accountability system is based on an agreed upon expectation of performance. At Southview School this expectation was the provincial standard, level 3 on the rating scale of 1 to 4. This was understood to be the collective expectations although some teachers (Grade 5/6 teacher, A02; Resource teacher, A02) felt it was an ambitious expectation for some students at the school. EQAO tests measured student performance against this external measure and the administrative position was that school performance should aim to meet the provincial benchmark.

We're always striving to meet Ministry expectations, always. That is the goal. It doesn't always happen because of the factors that are involved...Certainly as far as administration is concerned, and also myself, meeting the Ministry expectations is what we are striving for with our students. (Resource teacher, A02)

Mrs. Long felt that Southview students should meet those expectations, although her expectations might differ in other circumstances. She explained that the provincial standard was clearly delineated in the Ministry's Exemplars documents and these documents were a tool that teachers could use to understand performance criteria and assess student achievement. However, she also recognized that school expectations needed to be tempered by knowledge of the student population.

It all goes back to the teacher, who has to know who her students are and where the students are coming from. This particular school is not a multi-cultural school and this particular school has two parents at home. This

particular school does not have poverty and if I was a teacher or principal in a different school, I'd look at those exemplars as a standard [but] I would also have to look at who my kids are. And if some of my kids have a different background, then these exemplars may be the standard but do they meet the [suit my] expectations. (Principal, Ju02)

Mrs. Long understood the value of a standard but viewed its suitability for every student from a professional perspective. She has investigated her school community and worked with the stakeholders in developing collective expectations for school performance on the EQAO tests that were appropriate to her students. She had determined that the level of student achievement at Southview School should meet or exceed the provincial standard. The teaching staff understood the principal's perspective and accepted the standard as appropriate for the school, while acknowledging that some students may not achieve that level of performance. The Grades 5/6 teacher questioned the appropriateness of the provincial standard when "province wide results revealed 50% of Ontario students are failing year after year." However, this same teacher reported that "92% of his students had achieved levels 3 and 4", but that he was quite surprised sometimes. "Last year, I'm astonished at the fact, that there were three or four kids that I did not expect to get into the three level that did" (Grade 5/6 teacher, A02). The Grade 4/5 teacher (F02) commented that the curriculum and its rubric standard was a "sort of guideline" but that "in the end, the standards are what will be on the EQAO."

These comments revealed a general consensus that the provincial standard was ambitious but that most students should be able to attain it. Teachers (Grade 5/6 teacher, A02; Grade 7 teacher, F02) reported using the curriculum expectations and wording from the EQAO anchor books and exemplars to design projects and assessment rubrics. As the Grade 7 teacher (F02) explained, "it's not the expectations aligning with what I am doing but me having to align to the expectations."

Mrs. Long and her staff were consolidating a collective expectation for their school that was based on the provincial standard. Although she recognized that the provincial standard may not be an appropriate expectation for every student, the teachers (Grade 4/5 teacher, F02; Grade 5/6 teacher, A02; Grade 7 teacher, F02; Resource teacher, A02) at Southview School reported using the level 3 criteria when grading student performance. The principal promoted and monitored the use of the provincial standard as the appropriate school expectation.

Persons Who Judge School Performance

The fourth component of an internal accountability system requires that someone be responsible for judging school performance and this is a function of program review. At Southview School, the responsibility for this evaluation was found to be shared among many persons. The depth and breadth of involvement by district support staff, teachers, parents and school administrators in the program review process is described here.

Discussions of school performance on the EQAO tests and examination of all other relevant information were conducted by the principal and her administrative team with teachers, parents, community stakeholders and district support staff. The Committee attended a district SIP workshop where statistical analysis and school data were provided. Area superintendents and quality assurance personnel assisted with interpreting the data and consulted with the Southview administrative team.

The SIP Committee members spoke individually with the Grade 3 and 6 teachers about the EQAO results as part of the ongoing review of data. Afterward, the SIP Committee presented the EQAO results and identified trends in the data at a staff meeting so that “everyone is part of it” (Resource teacher, A02). The analysis of the

results was an administrative responsibility but interpretations of school performance and program effectiveness were discussed with all teachers. The principal explained her intention.

But we [administrative team] didn't make it critical of the staff, I don't want my teachers to feel stressed or pressured to teach solely to the test....I prefer the staff to sit and notice themselves and ask *What is going on here?* I can answer that but its always better if it comes from them, then they are a lot more motivated to make the change. (Principal, Ju02)

The analysis did not directly involve evaluating teacher performance but rather identifying what the trends in the data suggested about needed changes in the instructional program. In presenting the results to the staff, Mrs. Long enabled her teachers to notice trends in the data, flag gaps in achievement and realize where areas of strength existed. The Grade 7 teacher (M02) reported that "when we were presented with the results, they spoke for themselves, and we discussed among ourselves what we saw and that was the end of it".

Teachers' ownership for student achievement was described by the Resource Teacher (A02) who explained that the Grade 3 and 6 teachers sought her out when the scores arrived, "I mean that when the students are in your classroom and they're being tested, there is always more sense of responsibility." The Grade 4/5 teacher (M02) reported that, "The staff did talk about a school wide goal for improving the math. I don't know if it would be instruction per-say, but ...an overall improvement of the EQAO [math] results for the next year." The Resource teacher explained how follow-up to the program review was organized.

Results are shared with everybody in the staff meeting. When they meet as teams, they go over what each grade level can do to improve in this area. So there is really a push for all of the other grade level teachers to understand that it is not just the Grade 3 and the Grade 6 teacher who are responsible for the EQAO results, for getting the kids ready for the tests. What you do in

your year has a direct impact on how students do. So that's a huge push here and I think there is this understanding at the grade levels that, yes, we are getting them ready for EQAO. (Resource teacher, A02)

Teacher professionalism was respected by Mrs. Long, but this didn't mean that she was complacent in reviewing the provincial testing results with her staff. What was not apparent to the teachers was the input that the principal had received from other stakeholders in the community. The principal was party to some concerns expressed by the staff and parents from the feeder high school who reported that the intermediate students were "not ready for the Grade 9 math" and wanted "to ensure that the link was being made" with Southview School teachers (Principal, Ju02). In addition to this community dialogue, the Parent Council looked closely at the EQAO results and discussed them with the Principal. Mrs. Long found that with only the one exception she recalled, individual parents did not ask about the school's EQAO results. These findings suggest that district staff, the Parent Council and community leaders contributed to the analysis of school performance that informed the evaluation of the school instructional program.

At Southview School, the principal was the recipient of the data, the prime investigator of school performance, and the conduit for input from the community. In turn, she shared these observations with the administrative team, and then worked with all the staff to facilitate appropriate conclusions about the instructional program. From this collective evaluation process, Mrs. Long set goals and with her staff planned strategies to address programming needs.

The principal of Southview School was the active agent as she focussed staff attention on the steps needed to update resources, have workshops on new instructional strategies, and reorganize school procedures and practices. Collegial interaction was

used to ensure participation in the assessment process and commitment to the strategies put in place to enhance school instructional programs. The specific accountability instruments and strategies that guided teacher practice at Southview School are presented next.

Mechanisms to Guide / Monitor School Practice

The fifth component of an internal accountability system is the mechanisms that guide and monitor teachers' practices. Instruments and strategies reported at Southview School included external policy requirements (originating at the provincial and district levels) and strategies associated with the principal's strategic management plan. These are now reported in more detail.

External mechanisms. The external mechanisms at Southview School included the annual school report, the school improvement plan (SIP), the standardized report card, teacher performance appraisals (TPA), the Ministry target-setting exercise, and the district's school review audit.

Mrs. Long was required to update and distribute an annual school report. This external requirement profiled Southview School. It identified the school aims, outlined information on the school population and its community, described the instructional programs, and explained the special initiatives that parents and teachers were undertaking to support the school and its program. School EQAO performance was reported as a proportion (in %) in reading, writing and mathematics. Southview School Grade 3 and 6 results consistently surpassed these district and provincial percentages except in the areas of writing at the Grade 6 level. The report also outlined the goals and key activities of the SIP.

In describing the measures of student achievement and success, the Southview School Report (2001-2002) listed the ways that teachers assessed student achievement (rubrics, student portfolios, peer evaluation, test, quizzes, self-evaluation, student demonstrations and independent study projects). This list represented a formal statement to the public, and to the teaching staff for that matter, that Mrs. Long was progressive in her approach to assessment and encouraged students to be involved in the assessment process.

A second external mechanism was the *school improvement plan*. The SIP was one of Mrs. Long's main responsibilities because she coordinated SIP activities with district literacy, numeracy, and safe schools initiatives. She explained that the goals and outcomes "we developed last year, are working on this year and will be working on next year, go under those three headings" (Principal, Ju02). Once developed, this plan was shared with the Parent Council. It was also submitted to the District Quality Assurance Division.

Mrs. Long attended the district SIP workshop with her administrative team. At that time, she used the SIP manual to structure Southview's plan. Mrs. Long had implemented Southview's SIP in September based on information from last year's EQAO results and other data collected at that time. Her perspective on the value of continuous planning and the natural planning cycles of the school year are explained in the following comment.

I know the EQAO results don't come till the [late] fall and that might change your plan but it shouldn't. If we're looking at those literacy and numeracy goals, we're looking at them all the time. So it doesn't matter if your EQAO results got 90 or 50, you'll still be working on numeracy. You might want to look more specifically at something but the bottom line is that something has to be done for numeracy, regardless if your scores are 90s or 50s. (Principal, Ju02)

The SIP had been implemented in September after consultation with the teaching staff. The teachers proposed activities, while facilitated by the principal, were delegated to the primary, junior and intermediate teams to be carried out.

The Southview SIP listed three exit outcomes:

1. to improve students' levels of Ministry expectations in the writing area of the curriculum;
2. to maintain a safe and inviting school and to include members of the community within our school; and
3. to enhance teacher instruction and student achievement in mathematics.

(Southview SIP 2001-2002)

These goals aligned with the District's literacy and numeracy initiatives. For the purposes of determining how teacher practice was both guided and monitored at Southview School, the numeracy goal is examined more closely.

Seven specific activities were listed for enabling the mathematics outcome (SIP 2001-2002, p.2) and the responsibility for enacting them was assigned to the primary, junior and intermediate teams. The actions included: a) making an inventory of all math manipulatives, b) making an inventory of math textbooks and updating this resource (complete class sets), c) displaying math work throughout the school, d) using math computer programs, e) in-service teachers on math journals, computer programs and brain gym activities; and f) organizing math fun activities. The primary team also planned to establish math buddies for its students. This process was facilitated by Mrs. Long but involvement of the teaching teams in the early stages meant an implicit commitment to the SIP goals.

The EQAO required a SIP, the District provided guidelines and expertise, Mrs. Long and her administrative team framed the outcomes, and then teachers suggested

specific activities that they intended to enact. The means by which Mrs. Long further guided and monitored the implementation of the SIP is reported later under internal mechanisms.

A third external mechanism that guided teacher practice at Southview School was the *standardized report card*. The report card was organized so that each of the five mathematics strands that comprised the new Ontario curriculum was assessed and graded separately. The report card format was a visible means to monitor how teachers planned and assessed their instructional program over the school year and, in many cases, required teachers to reorganize their traditional way of both teaching and assessing mathematics. Central School District guided teacher practice by providing an assessment and grading handbook to adapt the provincial report card policy to a manageable level for teachers. The handbook required that each math strand be assessed twice during the school year. It suggested strategies for transcribing percentages to letter grades and advised on how to evaluate summative achievement from a series of recorded assessment marks into a single letter grade for the report card. By establishing these guidelines, teachers were not overwhelmed by attempting to evaluate all five strands each term, but neither were they permitted to maintain the status quo.

Mrs. Long gave feedback to her staff during the final stage of the report writing process and was responsible for monitoring that procedures were followed. The report card required that teachers know the new curriculum, its expectations, and how to use the supporting documents that clarified “characteristics of student work at each of the four levels of achievement” for each grade (The Ontario Curriculum- Exemplars, 2002, p. 4). As the Grade 5/6 teacher (A02) explained, “The same way my projects are taken from the expectations so I’m sure of covering the curriculum, my rubrics are taken from

the wording in the documents.” All these resources and supports assisted teachers in developing consistency in assessment and reporting.

A fourth external mechanism that the principal used to support accountability at Southview School was the *teacher performance appraisal (TPA)*. The principal was responsible for ensuring that TPAs were conducted regularly and she had completed seventeen in her first year as principal. The process involved formal observations and an interview. At the interview, Mrs. Long conducted *show and tell*, “I ask the teacher to bring everything that he or she wants to show me that is exciting in his classroom on assessment, or maybe something that he’s designed and is really proud of” (Principal, Ju02). In addition to interviews and reviews of instructional planning, Mrs. Long also visited the classrooms every day. This was something she did whether teachers were undergoing an appraisal or not. During class visits she reported making observations. These observations were a mechanism used to monitor teacher practice and contribute to the teacher appraisal’s based on the Standards of Practice outlined by the College of Teachers. She explained that classroom visits also helped her discuss professional development plans. It was interesting that teachers noted her interest in the students but did not comment on the professional reasons associated with her visits (Grade 4/5 teacher, M02; Grade 7 teacher, M02).

A fifth external mechanism newly implemented during the timeframe of this case study was the *Ministry target setting* exercise for literacy. This mandate was linked to student performance on the EQAO provincial tests. The target setting involved the primary teaching staff and is described using Mrs. Long’s own words.

This year we had to add that target setting sheet where we predicted where the students’ growth would be in three years time. The target-setting sheet was really interesting. We took all our kindergarten, grade one and two

students and noted where they are now, where they are achieving against the Ministry expectations, and what we need to do to make sure that in grade 3 they meet expectations. ...Then I showed all the staff the targets. I'll show this to the primary and junior teams and then, usually, the Grade 3 and 6 teachers are the ones who want to sit one-on-one and go through it more carefully because it affects them personally, right. It was their students and they wanted to look at certain students' scores. (Principal, Ju02)

These targets were submitted to the District and reported to the Ministry. This mechanism was external but it was the responsibility of the primary teachers to determine the literacy target (achievement level) based on their knowledge of the individual students at the school. The target setting activity required teacher expertise and professional commitment. This obligation to meet the achievement target set was a powerful motivation for professional reasons, but also because the success of the instructional program was monitored and measured by the EQAO test scores. In contrast, ownership for students not achieving the school-developed target would involve personal professional sanction at the least. It was anticipated that a numeracy target setting exercise may follow in the near future (Quality Assurance coordinator, Au02) but no such requirement has yet been reported. No junior or intermediate target setting was mentioned and the full impact of this exercise, what it meant to staff, and its impact on internal accountability could not be ascertained fully within the timeframe of this study.

The *School review audit* was the sixth external mechanism reported at Southview School. The review process was a district-sponsored program that used the principles of effective schools theory (School Review guidelines, 2000) to monitor school organizational management and program implementation. Mrs. Long had conducted the review the year previous to this study. Under the supervision of the superintendent, she consulted with the staff, students and parents to determine: a) school expectations for

learning and accountability, b) the nature of the school organization, c) training and staffing needs, d) school and community partnerships, e) safe school status, and f) use of school facilities. This review enabled an analysis of school organizational strengths and needs, and from this analysis a strategic management plan was formulated. The school review process was originally supervised by the district superintendent, who then monitored the strategic plan the following year.

Internal Mechanisms. The six internal accountability mechanisms reported at Southview School were closely associated with the principal's strategic management plan. These included staff reorganization, annual personal growth plans, new instructional materials, working staff meetings, long-range plans, and timetabling. These mechanisms are reported in further detail immediately below.

The *principal's strategic plan* involved management tools and strategies to both guide and monitor teachers' programming and professional interactions. It was developed as a follow-up to the School review audit. The superintendent touched base periodically on the progress of the strategic plan and Mrs. Long discussed strategies and sought advice from other administrative colleagues as well. A clear hierarchy of responsibility was established at Southview School with the strategic plan being the principal's responsibility and facilitating the SIP one of its main activities. Mrs. Long managed organizational routines and created venues for professional dialogue as part of her plan.

One of the first mechanisms that Mrs. Long employed was to address the issue of staff renewal through *team reorganization*. She orchestrated a change in teaching assignments when a number of retirements the previous summer occurred. This provided her with an opportunity to reassign some senior teachers to a different grade level. She

reassigned staff so that two of the teams (primary, intermediate) would have a better balance between experienced and beginning teachers. In reviewing teacher practices, Mrs. Long made changes that “weren’t necessarily popular at the time” but were based on her knowledge of teachers’ capabilities and strengths. She explained that “if I know that a teacher will be better here [in Grade 3] because she will, or he will, do a better job serving that group of students, then I’ll do that.” This mechanism guided teachers towards areas suited to their skills. Mrs. Long had wanted a new dynamic at the intermediate level and found that reassigning teachers had “rejuvenated the teaching team” by bringing in teachers with new perspectives. In the process, the new groupings also led to the formation of “an exceptional primary team” who could “take the lead in how they were going to proceed” when an experienced intermediate teacher was reassigned there (Principal, Ju02). As she explained it, senior teachers brought time management and behaviour management expertise to the teams, while the newer staff brought in new ideas and generated a willingness to “try another way.” Mrs. Long was in the process of finalizing the same reorganization at the junior level, as two more teachers were retiring at the end of the school year.

In moving forward her strategic plan, Mrs. Long used a second mechanism to guide teachers’ practices. She requested that all teacher identify one goal in their *annual professional growth plan* that was “related to one of the three points” in the SIP. These professional obligations were easily monitored by classroom visits and teacher reflection in formal (TPA) and less formal (annual professional plan) meetings.

Mrs. Long set aside funds for workshops and school visitations to support the growth plans. She was surprised however, that when reviewing the growth plans with her staff, some teachers “felt that they hadn’t had the opportunity” to attend a workshop

during the school day (Principal, Ju02). She encouraged staff to go to workshops, but she expected teachers to seek out opportunities for themselves. Many teachers had not taken advantage of the funding. Comments from teachers (Grade 4/5 teacher M02; Resource teacher, A02) suggest that this was so because many workshops were after school and frequently at a distance. In her remaining year-end meetings, Mrs. Long anticipated chagrin by a few teachers as she fulfilled her responsibility as instructional leader and met to discuss progress on professional goals. How the impact of this professional discourse played out at Southview School would be an interesting extension to the study.

A third internal mechanism that guided teacher practices at Southview School was the removal of out-dated mathematics textbooks and the purchase of *new instructional materials*. As instructional leader, Mrs. Long supervised the removal of outdated math books because, in her words, “We did an inventory and got rid of old books...physically take [took] them out of the building so that they weren’t accessible anymore” (Principal, Ju02). There were better resources, especially for the English program, and the Immersion teachers had been badly in need of new math books.

Mrs. Long reported that teachers had been comfortable using the old materials and found the change difficult because the newer math resources were integrated in format and involved performance-based activities. Teachers made the following comments about the suitability of the new books. “The Interactions textbook is basically all I’ve been given. It does not cover the curriculum totally. It has some good things in it, but is very uneven” (Grade 5/6 teacher, A02) and “usually the culminating test is a mixture of both, so it will be the Quest 2000 final test and I’ll add some of my own questions” to cover the needed outcome assessment (Grade 4/5 teacher, F02). The Grade

7 teacher (F02) did not have a class set of texts and found that he had to “use *Math Power*, *Houghton-Mifflin Math*, *Math in Context* and another document that I bought” to plan lessons. The intermediate math program was reported to be under resourced.

It is interesting to note that these comments were a fair criticism of the textbooks, but were also a valid criticism of the previous textbooks that had been even less aligned with the curriculum expectations. The transition to new texts and instructional materials created a period of adjustment as staff became familiar with the new materials, cross checked the lessons with the new curriculum, and sought additional resources to meet the diverse expectations. In this instance, change was triggered from feedback on student performance.

The new instructional materials facilitated a different instructional approach and the use of performance assessments. Teachers found that in using the curriculum guide they had to find suitable math activities from the textbooks, anchor books, and the newly arrived exemplars. Curriculum-linked activities from the Ministry websites were also used and integrated units of study were accessed via the Ontario Curriculum Planner. If the teachers had been actively seeking resources that assessed ‘performance’ as well as ‘product’ earlier on, the change in textbooks would have been welcomed. As the situation was reported, the change in instructional resources necessitated a change in instructional (and assessment) practice. This reminds one of the classic question: which comes, first the chicken or the egg? Removing the old textbooks was a leadership intervention that guided all teachers to use the new resources.

Another administrative intervention, that followed the inventory of math manipulatives, was the centralization of these materials. This enabled sharing of the math manipulatives, as Mrs. Long pointed out:

In reorganizing the manipulatives, we noticed that we had tons of stuff within the school but it had been kept in one or two classrooms, so was not accessible to all classes. These are now stored together and one teacher is responsible for it, checking it out and in, so that everyone can use it. So in math right now, as far as available resources, we are well equipped all the way from primary to intermediate. (Principal, Ju02)

The manipulatives were originally viewed as primary resources even though some junior teachers knew they could be borrowed (Grade 4/5 teacher, M02). With these resources more accessible, teachers from the junior and intermediate classes were encouraged to use them and incorporate them in their lessons. EQAO performance tasks often required familiarity with such resources and Mrs. Long used the inventory to raise awareness of what was available and then saw that their use was monitored.

Mrs. Long made changes in school operations to foster professional learning; the fourth one was *working staff meetings*. Mrs. Long reported that information that could be shared by e-mail was no longer discussed at staff meetings. She preferred to work with her staff, rather than merely direct them, so the monthly staff meeting was for planning and professional learning. She encouraged team meetings because they were opportunities for collegial dialogue and discourse. She attended all meetings so that she could support professional learning and keep improvement plans on track by promoting new ideas and monitoring progress on SIP goals. At these meetings, the principal encouraged teacher leaders to share 'best practices' but didn't insist on teachers making immediate changes in what they did. She recognized teacher innovation and encouraged teachers to take pride in sharing effective strategies as the following quote reveals.

We have an exceptional teacher at the intermediate level who does math journals. He did a presentation to the whole staff because no one really knew what math journals were. He brought in samples, and then the staff talked within their groups and discussed how this strategy could work for the primary classes, the junior classes, [and *sic*] how it could work for the intermediates. Yes, I would love for all my teachers to be doing it, but I don't

push them to the point where they feel they can't take on any more. It's too much new information, but I do want to introduce them to new strategies out there. Then I want them to use their own teaching style and know that they can take ideas from others. (Principal, Ju02)

A fifth strategy traditionally used at Southview School to monitor teachers' practice was *long range plans*. Mrs. Long reported that she was not interested in the massive binders that teachers prepared because she was certain that "they weren't using it...It was for show, not necessarily what they were doing"(Ju02). She designed a *one-page long range plan template* and asked her staff to use these at the May staff meeting where planning for the next year had begun. She explained:

We actually did long range plans at our staff meeting for next year. If you take mathematics, I didn't want the nitty gritty, I wanted to know which strands the teachers are covering and when they are covering them. For example, which term, how many times do you think you are going to cover them....all that goes in that one sheet format. (Principal, Ju02)

Mrs. Long explained that she wanted a simple format for plans so that she and her teachers could sit down and discuss progress easily. She wanted to have her staff review what they had done in a term, where they were now and what they needed in order to complete their instructional program. She wanted to be more closely involved in monitoring plan completion and wanted her teachers focussing on getting the job done as opposed to creating a vast collection of lesson plans that were not meaningful (Principal, Ju02). Mrs. Long believed in being a coach to her staff and saw collegial dialogue and ongoing monitoring an important responsibility. This is reflective of formative assessment except that it was a focused on providing feedback and support on the process of teaching the instructional program.

A sixth internal mechanism used to guide teachers' practice at Southview School was *timetabling*. Class timetables were often organized to match the availability of

resource support for special needs students. This meant that certain periods during the day lost valuable instructional time to morning and end of day routines. Of particular concern was the reduced time for intermediate math, which was frequently slotted the last period of the school day when students were tired and valuable instructional minutes were lost because students organized homework and readied to catch the bus before the bell (Grade 7 teacher, M02). To address this concern and to take into account the high school feedback about student preparedness for the Grade 9 math program, Mrs. Long rescheduled all the Grade 7/8 math classes to second period because it happened to have an additional ten minutes. She hoped this common time would initiate team planning and collaboration. She shared her rationale:

Grade 7 and 8 is really loaded with a lot of curriculum that kids aren't getting through. So we thought that extending time, having it all at the same time, and having smaller groups because the SELC [Special Education Learning Centre] children in the Grade 7 and 8 classes will be removed for special programming [would help]. So classes will be sixteen, seventeen, twenty students instead of 33. The other students can have more direct teacher time in the smaller classes. The teachers can instruct in different styles, and if we need to regroup kids for a certain time, we will be able to do that for math. (Principal, Ju02)

Mrs. Long responded to suggestions from teachers who wanted more team collaboration and who recognized that “with the exception of our divisional meetings, it is always informal. There are no structured meetings or times, if there was, it would be phenomenal” (Grade 7 teacher, M02). As principal, Mrs. Long owned her role as instructional leader and responded to input from staff and community members, and in doing so, managed the changes to the school operations.

These mechanisms relied on leadership and professionalism at the school level. Many involved feedback to the teachers and principal via discussions and consultations between professionals, colleagues and parent stakeholders. This feedback was formative

because it informed school decision-making, identified real concerns, and led to site-specific adaptations and changes focussed on instructional improvement.

Southview School's internal accountability system possessed aims stated in the School Report and the SIP. It possessed a standard of achievement and information on school performance to measure of the effectiveness of school programming. Mechanisms to guide and monitor teachers' practices were in place. It must be noted that teachers did not readily identify many mechanisms that were being used to guide and monitor their teaching practices. They seemed unaware of the strategic plan and the next steps that were to be implemented. The teachers reported that Mrs. Long consulted with them regularly (Grade 4/5, M02; Grade 5/6, A02; Grade 7, M02) and was very supportive of their efforts to use new resources, but Mrs. Long reported that the initiation and management of the school renewal activities was a principal-centered responsibility. Mrs. Long's concerns (J02) that her staff was under a lot of stress implementing new curricula, familiarizing themselves with new instructional materials, and modifying instructional units to incorporate new teaching and assessment practices. She wanted their energies focused on the students and the teaching and this is what she consistently communicated.

Factors Impacting on Southview School's Internal Accountability System

This section addresses question 1b of the study and describes the policy to practice factors most closely linked to external policy implementation and change theory. For the purpose of this study, the two main factors reported here are *policy requirements* and mobilization of *supportive actors*. See Figure H7 for a synopsis. The first factor is elaborated as specific elements inherent in EQAO performance testing program; while the second factor is interpreted as a function of school leadership. Since

mobilizing supportive actors was linked to political issues in the context of the study, this element is briefly reported as well in this section.

EQAO Performance Tests

The EQAO tests influenced the Southview School internal accountability system in three ways: firstly, it provided an external standardized measure of school performance; secondly, it modeled good instructional (assessment) activities; and thirdly, it captured the attention of teachers and members of the public. Specific impacts for these areas are reported next.

The EQAO test results were an external measure of Southview School performance against the provincial curriculum expectations. One impact on internal accountability was that they provided an external measuring stick that Southview School administrators and teachers could use to identify trends in the scores, flag strengths and weaknesses in the instructional program, and compare school performance from year to year. The EQAO scores, along with other information, formed the basis for the program review. A second impact of the EQAO tests was their role in the Ministry target-setting exercise. Teacher goal setting and the effectiveness of the primary instructional program was to be measured against student performance on the Grade 3 EQAO reading test. This focused school accountability on developing individual student learning plans to approach “meeting Ministry expectations” (Resource teacher, A02). This exercise involved a collaborative effort by all primary teachers and distributed responsibility for student progress to teachers at each grade level.

A third impact of the EQAO tests was increased awareness and commitment to addressing all curricular expectations (i.e., all five mathematical strands) at each grade. The Southview School staff selected new math textbooks and resources that were “in

line with the provincial guidelines, the provincial curriculum” (Resource teacher, A02).

As one junior teacher explained,

When I began [teaching] we were using this approach already. From other teachers’ perspectives, the ones who are teaching in the primary grades now, it’s fine now but they’re kind of starting it with grade 1. The teachers that are teaching intermediate level, I think they are still seeing the problems with it because it started being implemented after their students had already gone through a number of years with the previous curriculum. (Grade 4/5, M02)

The Grade 5/6 teacher characterizes the change from the junior teachers’ perspective as they began to use the new instructional materials.

Resistance, anxiety, the old way is better, the fundamentals, but then you try the new way and you figure out, while you do lose some of the fundamentals, nothing is perfect. But if you supplement it [the approach in the new textbooks] with the traditional approach and concentrate on the fundamentals as best you can, you get by. So, for forcing teachers to cover the new curriculum, it is good, that is the only merit I see. It’s a big one though. (Grade 5/6 teacher, A02)

The principal encouraged school leaders to acquaint themselves with strategies known as *best practices*. The Grade 7 teacher realized that he needed to teach his students to reason and communicate in mathematics and started math journaling with his students.

My most recent change has been the journal. I heard about this idea and after seeing the EQAO tests, it just reinforced the point that the students have to be able to explain their thinking, be able to show their work. The students have to show their answers in a multitude of different ways. The most logical way to do that is through the journal. (Grade 7 teacher, F02)

Journaling was part of the SIP and this teacher prepared a short workshop presentation on math journaling for a Southview School staff meeting.

The EQAO tests triggered a fourth impact on the internal accountability system; they raised teacher awareness of new assessment practices. The tests themselves provided samples of performance tasks that could be used in classrooms since the tests changed each year. Teachers used word problems modeled from previous EQAO tests.

I looked at previous EQAO testing material and picked some questions to do with my students. I made sure that they were aware of the wording of the questions and so on. When students have problems to solve, the wording is quite difficult for them and I made sure that the students got used to that. I did practice questions before they wrote last year's test. (Grade 4/5 teacher, M02)

We are redoing old tests since they are now available. I take at least four to six weeks doing it. Why not? Prepared curriculum. It's exactly what we're supposed to be teaching and I use this curriculum. I found a way to get something I can teach. (Grade 5/6 teacher, A02)

Initially, teachers found that they *taught to the test* to familiarize students with the wording of questions and to practice showing their reasoning in different ways. However, the Grade 5/6 teacher (A02) explained that he also used EQAO tests and anchor books as a source of assessment tasks for his own mathematics program. The benefit was ready-made tasks from every curriculum strand with assessment guides containing evaluative criteria.

A fifth impact of the EQAO tests on internal accountability was related to the opportunity for teachers to compare and contrast their expectations for students' achievement with those of the trained markers, as the following comments elucidate.

More or less [EQAO performance was within expected range] but with the exception of very specific areas. Some students had scored a level 2 when I thought they were a level 4, but overall, in terms of their performance and looking at the overall grades, it was pretty consistent with what was happening in the classroom (Grade 7 teacher, M02).

Yes, definitely, the scores confirmed my own assessment of the students. And I was surprised too because I thought the scores would be different, you-know. The students' results would be different than what they were doing in the classroom, but the student results were similar to the classroom work being done this year (Grade 4/5 teacher, M02).

The EQAO results were reviewed by the receiving teacher and by the former teacher at Southview School. Teachers reported that the performance scores confirmed their expectations in general, although the teaching staff and principal felt that these test

scores presented only a partial picture of student achievement. Three reasons for this perspective were given: a) one was the *time limit* because conscientious students included a lot of detail in each question and may not have completed all the test items. (Grade 6 teacher, A02); b) a second was the *frustration* felt by students when teachers were not able to clarify explanations or questions (Grade 4 teacher, M02); and c) teacher concerns for special needs students who *emotionally cannot handle it* (Resource teacher, A02).

The anchor books and the Ministry *exemplar books* were used by the junior and intermediate teachers to study grade-appropriate performance tasks and to visualize assessment criteria for each of the four levels of achievement. The rubrics and comments modeled “where the child should be in terms of skills and knowledge” (Resource teacher, A02) and the samples of graded student work were accompanied by “descriptions of what each level is” (Grade 4/5 teacher, F02).

The EQAO tests were reported to impact on the internal accountability system at Southview School in a sixth way, school reputation. Since the aggregated school and district scores were published on the EQAO website, in local newspapers, and distributed to parents, the scores were open to public scrutiny and comparison with other schools. The Resource teacher (A02) reported this led to “a lot more stress on teachers because of that. Certainly because the community at large looks in the newspaper and sees where your school ranks.”

Public reporting brought public opinion and input to school accountability. It triggered professional discussions. This accountability dynamic, while not viewed as an evaluation of teaching staff, was an evaluation of the instructional program, and it did contribute to school reputation and professional esteem.

The EQAO tests influenced Southview School in many ways. These tests received a lot of teacher attention because teachers were interested in their design, the performance scores and the feedback they provided on provincial expectations (Resource teacher, A02). The internal accountability system was influenced by these tests because they: a) provided quality examples of instructional activities, b) revealed strengths and weaknesses in the instructional program, c) assessed all five mathematics strands, and d) tracked the effectiveness of improvement plans.

Leadership

Leadership was a second factor that impacted on Southview School's internal accountability system. Mrs. Long's leadership was important for two reasons. The first reason is that she supported collegial dialogue and professional ownership for a quality instructional program. The second reason is she assumed responsibility for organizational changes to maximize the use of the available human and technical resources.

Mrs. Long guided her staff by consulting on school improvement planning and by focussing on strategic priorities so that expectations remained reasonable and the stress of change on teachers kept to a manageable pace. She recognized the value of collegial dialogue to sustain motivation and cope with transitions. Mrs. Long supported system restructuring but in her own words recognized that: "It was too fast, but it needed to be done" (Ju02). She knew that changing instructional materials was a significant first step and sought to maximize collegial support because "everyone has to improve. It doesn't matter who you are." She appreciated that schools had to be accountable.

Mrs. Long assessed the capabilities of her teachers and encouraged teaching teams to stay focussed on SIP goals by participating in team meetings and working with team leaders.

I have an exceptional primary team at this school, exceptional. I have a good junior team, it will be better next year because there have been changes [planned], and my intermediate team is just being rebuilt because last year five teachers retired. So they're [the teams] all at different stages and I think, as their principal, I need to understand where they are at, and how much I can bring them along. (Principal, Ju02)

When Mrs. Long conducted the School review audit, she assessed school operations, inventoried resources, prepared a needs list, and consulted with stakeholders about the school, its mission and its goals. In explaining her strategic plan, she reported many organizational routines and practices had been in place for a very long time and that she felt changes were needed to support new expectations. Her strategic plan enacted in September, had already addressed staffing issues the previous June when new teachers were hired.

The aims of the internal accountability system were facilitated by the three stages of Mrs. Long's strategic plan, which she described as a "work in progress" (Principal, Ju02). Each stage involved administrative decisions that resulted in changes to organizational procedures and in opportunities for professional learning. The first stage addressed the issue of resources. The second stage was school-based professional development structured through collegial dialogue and discourse. The third stage, which Mrs. Long was actively organizing at the time of the study, was to restructure timetabling and teacher monitoring procedures.

Mrs. Long was sensitive to the demand these changes made on teachers' time. She recognized the stress this created and reported that it was important to

“always keep reminding ourselves that we are working on literacy, numeracy and safe schools, and asking if it falls under those three goals. No. Then let’s not do it, it’s too much” (Principal, Ju02). This may explain why teacher involvement in the School review and SIP processes was more consultative than dynamic in term of the decisions and organizational changes.

Mrs. Long mentored *professional learning* through the annual growth plans. Although the teachers at Southview School indicated an interest in more professional development, the Resource teacher (A02) reported there was a stated preference for “people coming into the school and giving workshops and in-services.” The principal had reservations about this approach as she explained:

I think teachers need to meet together...Yes, it’s hard to find time to do that, but you have to have teachers who are dedicated to make the most of that time. If I was to release all my primary teachers...I know they would be working on curriculum, but if I offered that to someone else, unless that someone was on top of their own professional needs, it would be a wasted half-day. (Principal, Ju02)

Instead, Mrs. Long structured the monthly staff and team meetings differently and changed their purpose.

My last few staff meetings in the third term have been working staff meetings. That was another part of my personal plan when I arrived here last year. Staff meetings were just information, which today, I think can be shared in other ways; ...I think the meeting should be used to discuss teaching and the methods we are using. Talking about teaching strategies is a lot more important. So what we have been doing now is ‘teaching.’ I call them working staff meetings. (Principal, Ju02)

Mrs. Long encouraged teachers to share ideas and resources at staff meetings and this indicates how professional learning through *collegial dialogue* was being facilitated at Southview School.

Another impact on internal accountability resulting from Mrs. Long's leadership was the collaborative planning she initiated. Team planning was common practice among the primary teachers and, as principal, Mrs. Long formally structured planning sessions for the junior and intermediate teams at the working staff meetings. "I would like the teachers to sit together and plan units around mathematics and that will be an expectation for next year. I would do it by primary, junior and intermediate groups. My last few staff meetings, in third term, have been working meetings" (Principal, Ju02). This strategy was intended to develop common expectations and to support risk-taking as new programs are developed and best practices introduced. This strategy established individual accountability to school aims and expectations. At the time of the study, it was premature to ascertain if this led to meaningful collaboration or contrived collegiality.

As principal, Mrs. Long did not want her teachers to lose ownership for how they taught but to reflect on options that have worked for others and reconsider their current practices in light of professional dialogue and discourse. Her perspective of the teachers' responsibility was to provide an instructional program that addressed all needs of the student, and this included addressing matters of class management. She spoke to this point when discussing how she envisioned her leadership role in the school changing from one of disciplinarian to one of instructional leader.

I definitely want a session on disciplining children because we all have a lot of different styles in this school. Not that we all have to be the same, but, there are different strategies that the staff could use. There are things [responsibilities] that I call classroom management. For example, if a child is chewing gum, he shouldn't be sent to the office and certain staff members need to understand this and ask themselves, "What could I do in my classroom to make it work?" By sharing ideas about what works for you, staff can think about what works for them [*sic me*]...I see a child going to the

office when its beyond anything that a teacher can handle in the classroom.
(Principal, Ju02)

Mrs. Long defined her role as instructional leader. She explained that “it’s really the leaders of the school, the principal and vice principal, who have to know if you can offer that time for planning to them and whether it will be beneficial.” She also explained that at the present time in the early stages of her strategic plan that her guidance was needed, “because if I’m there, then it will get done” (Principal, Ju02).

Mrs. Long recognized Southview teachers were at different stages in implementing the new curriculum and using performance tasks and assessment strategies. She respected this situation and set her expectations accordingly. Working staff meetings and classroom visits monitored classroom dynamics but these actions also addressed issues of isolation (Grade 5/6 teacher, A02; Grade 7 teacher, M02) and encouraged teachers to “ask for input or anything we need to add about what is going on with our programs, and that’s where she is quite helpful” (Grade 4/5 teacher, M02).

In addition to forward thinking, Mrs. Long also recognized the importance of celebrating and giving *recognition* where it was due. Just as the teachers organized assemblies for celebrating student learning as part of the SIP, she was organizing a celebration at year-end for her teachers (Principal, Ju02). She saw this as an opportunity to recognize “everything that was so great that we have done this year” and planned to display student work, teacher achievement, and school projects in the library. She was optimistic that this recognition would encourage the junior team in particular and motivate others to take the next step.

Leadership impacted the internal accountability system at Southview School because it established priorities, modeled high expectations, paced the introduction of

organizational changes and mobilized teacher collaboration and networking. Mrs. Long, as instructional leader guided and supported her staff. She planned for celebrations of teacher achievements and sought to motivate with professional recognition. Her leadership was an enabling factor to the internal accountability system.

Political Constraints

Another element related to the second factor that impacted the internal accountability system at Southview School was a political issue related to changes in governance of teacher certification. The College of Teachers introduced new requirements for professional re-certification and this resulted in a dispute with the provincial teachers' unions. The unions lobbied on the teachers' behalf and supported teacher ownership for personal professional development goals. As a result, teachers were asked to boycott College sponsored PLP (professional learning programs) courses and all professional activities associated with the EQAO not directly related to the administration of the provincial tests in the schools (i.e., the summer EQAO test marking sessions). Teacher unions encouraged *personal* professional development and qualifications *upgrading* through university courses but advised against participation in any program identified with PLP credits.

As a result of this political discord, teachers (Grade 7 teacher, M02; Grade 4/5 teacher, M02) were not participating in the EQAO summer marking session or being involved in EQAO test development activities. These valuable learning opportunities were boycotted by Southview School teaching staff and represented a lost avenue of influence to inform teacher practice. The EQAO tests and leadership were both factors that enabled the internal accountability system at Southview School. The political climate associated with teacher certification appeared to hinder it.

The Relationship with the External Accountability System

This section reports on the relationship between the internal and external accountability systems and speaks to the guiding question 1c. In examining this relationship, this section describes the degree of alignment reported among the aims and combined strategies operant between each system. The compatibility of the systems is summarized for each of the five components in Figure H8. This section addresses the intermediary role of the district within the relationship. For the purposes of this study of internal accountability systems, only instruments and strategies directly associated with the EQAO accountability program are described. It is understood that alignment between related restructuring frameworks (curricula, professional standards, fiscal resources) are not addressed here.

Compatibility of Aims of Accountability

The relationship between the internal accountability system at Southview School and the external system operating in the school's nested context is described here in terms of the stated aims and the level of accountability that the school was expected to provide for meeting expectations. The *aims* of accountability for the school, district and provincial system levels all focussed on two purposes: instructional improvement and on public accountability through transparent reporting.

At the *EQAO level* (*Provincial Report on Achievement*, 1997), the aims of assuring greater accountability and improved instructional programming are principally addressed through the large-scale performance tests that provide school performance scores. The scores represent a measure of performance against the provincial standard of

achievement. Districts and schools are required to report district and school aggregated averages to parents and the public. The EQAO, as the provincial authority, provided feedback on the DIP and published recommendations for school activities. The EQAO position was that schools and districts should employ the EQAO results as part of an extensive collection of information on performance to enable informed improvement planning. It supported transparent reporting but not ranking of schools and districts based on EQAO scores.

At the district level, the Central School District (Strategic Plan, 2000) aims to support student learning and guide school improvement efforts through operational planning, informed policy and budget decisions, and focussed district literacy and numeracy initiatives. The District consulted with school and community leaders to first define and then act on these aims through system-wide reviews (audits) and system improvements plans. Feedback on the DIP is received from the EQAO in the form of a letter with some suggestions for future consideration (District coordinator, Au02) and, in its turn, the District provides guidance to Southview School through the School review audit (Principal's strategic plan) and the SIP process (district initiatives and supports).

The aims of the Central District plans are enacted through its organizational support to school reporting and improvement planning. The accepted standard of achievement for grading student performance is the provincial standard unless students have Individual Educational Plans [IEPs]. District policy affirms that EQAO scores are part of the data used to review school performance and enlisted the expertise of the quality assurance officer interpreting and reporting of school performance.

The aims of the internal accountability system at Southview School focussed on the school instructional program through its program review (evaluation) and

improvement planning (SIP). Inherent in the program review, was reporting and consulting with parents and members of the public through Parent Council presentations and the distribution of the annual School report. The principal, working with her administrative team, interprets and communicates the EQAO results to Southview teachers prior to reflecting on program effectiveness and setting the goals for the SIP. The school principal manages the communication among district personnel, parents and community stakeholders. She is responsible for reporting and explaining the EQAO results and justifying the SIP. The principal affirms that the applicable standard of achievement for assessing students is the provincial standard and mentors Southview staff in this regard.

At all three system levels, transparency in reporting EQAO results and publishing improvement plans serves to meet the dual purposes of Ontario's professional practice approach to accountability. The level of the accountability involves obligatory reporting, program review and justification of improvement plans.

Compatibility of Information on School Performance

All three levels of the accountability framework supported the use of multiple sources of data to measure school performance. The EQAO supplied a standardized external measure based on a performance test and gathered other indicator data from surveys. To this information, standardized test results and information on socio-economic factors were supplied by the District. Southview School added report cards grades and input from IEP documents, teachers and parents. This collection was extensive and was compatible with the stated aims.

Compatibility of Standard of Achievement

The provincial standard of achievement was modeled by the EQAO tests and results. It was used as the benchmark for measuring the school performance at Southview School. It enabled comparisons from year to year and was one factor used in reviewing the effectiveness of the instructional program by the Southview School principal. In terms of collective school expectation, the provincial standard was reported as an appropriate measure of student achievement.

Compatibility of Persons Judging School Performance

District and school consultations with colleagues, parents and members of the community provided input into the program review process. At the District level, consultations with Board members and the general public extended the breadth of the accounting process. At Southview School, the principal consulted with colleagues, teachers, parents and members of the community as she communicated and interpreted the information on school performance and outlined the direction of the school improvement plans. The depth and breadth of involvement in the program review process was extensive.

Compatibility of Accountability Mechanisms

The instruments and strategies that guide and monitor system accountability are presented in *Figure H8*. This section examines the degree of compatibility between these instruments to determine if Southview School is able to comply with external requirements while balancing its responsibility to best meet the needs of its student population.

EQAO instruments and strategies. The EQAO conducts large-scale performance tests and collects information on the non-school factors related to student learning both inside and outside the school. The EQAO results are sent to the district and the school.

Provincial, district and school results are posted on the EQAO website and districts and school are required to report this information to the parents. The data are not intended as the sole tool for evaluating school programs nor is it intended to rank schools (EQAO, *Provincial Report on Achievement*, 1998). However, this is no guarantee as to how public interest groups will use the information.

The EQAO Assessment Officer (N02) interviewed for this study explained that interpreting the EQAO data was the responsibility of the director of education or his appointee at the district level and the principal at the school level. The EQAO Guide to School and Board Planning suggested a process that involved: group analysis of all the data collected, consideration of contextual information, improvement planning and communicating about the process with school and community stakeholders. In the EQAO officer's words, "The Guide is that, a guide, because the Boards have much more detailed processes" (N02). EQAO officers review each DIP against some of the strategies and give feedback to the district in the form of a short letter. This feedback is generally brief and includes "a commentary on what are their strengths and what are the areas to think about next time" (District coordinator, Au02; EQAO officer, N02). In describing the purpose for DIP, the EQAO representative explained:

It wasn't very long ago that Boards did not universally believe that they should write an improvement plan. In the first year, when we called for plans, there were some boards who were not accustomed to doing that at all and it was a bit of a struggle for them. Whereas, if I look at this past year when Boards were asked to hand in their plans, some of them sent them in immediately. In other words, they are into a planning cycle where that was not commonly or consistently found across the province previously. Also, the

quality of the plans, for example how detailed and articulate they are, has changed quite a bit...At this point, we are talking about a paper plan. We don't monitor whether or not the paper plan has been implemented. That is the responsibility of the Board." (N02)

Recommendations are meant to guide the process, but responsibility for enabling school improvement is delegated to the school district.

Central District instruments and strategies. The Central School District complied with the EQAO policies and requirements. It reported district and school EQAO results and developed a DIP. However, this plan was developed but "possibly not in the way you would expect" (District coordinator, Au02). The District developed a three year system-wide strategic plan that led to a series of system improvement plans (see Chapter 2).

The District's System Improvement Plan was more extensive than what EQAO required and this meant that Central District incorporated the goals of the provincial DIP as part of its own, more extensive System Improvement Plan.

They have a format of their own which is quite different from ours. So I prepare a summary of what we have done in their format. I also attach what I think are the key action plans from all our divisions (staff development, program, quality assurance) that relate to the kinds of things that they are talking about. We also now have district wide system improvement plans and these are developed quite differently and incorporate relationships with a lot of other factors, but sometimes these may be relevant to what they [EQAO] are doing as well. (District coordinator, Au02)

The System Improvement Plans required school review audits and strategic plans for superintendent review. The audit assessed six quality indicators arising from effective schools research: expectations for learning and accountability, learning organization, professional leadership, community-school partnerships, safe schools, and

school facilities. It was from this school review audit that strategic plans were developed.

To support school improvement, four manuals were developed in-house and distributed to schools. These manuals provided district guidelines for managing school audit processes, SIPs, interpreting statistics, and assessing and grading student achievement.

Southview School instruments and strategies. The principal's *strategic plan*, although less visible to the staff as a whole than the *School report* and *SIP*, involved decisions related to staff reorganization, structuring professional collaboration, guiding professional growth plans, and promoting professional development. Part of this strategic plan (Principal, Ju02) was the development of the SIP. Mrs. Long consulted with school stakeholders and worked closely with her superintendent. She also sought advice from a network of principals.

The *School review audit* and *EQAO testing* were external requirements that structured and guided evaluations of organizational management and instructional programs. These instruments helped identify organizational strengths and weaknesses. The EQAO results were an external measuring stick to inform on schools about programming effectiveness. The annual program review and SIP guided the changes in the instructional programs (teachers' practices) and were a means to ensure ongoing accountability.

These external requirements and internal mechanisms provided structured the school-based accountability but had sufficient discretion to allow Southview School priorities and needs to be addressed. The principal, Mrs. Long, assumed much of the responsibility for school-based decisions and delegated this responsibility when she saw

fit. In turn, teachers were responsible for their individual classroom programs. However, the principal, as instructional leader, sought to support teacher learning and monitor program development through internal mechanisms. This reflected a balance between external oversight and internal responsibility.

This balance suggests compatibility between the external and internal strategies. The current accountability agenda was something that Mrs. Long (J02) felt “needed to be done.” Curriculum renewal, changes in assessment, and accountability reporting were noticeable at Southview School. These represented a positive response to the EQAO agenda. In capturing the relationship between the internal and external systems, the EQAO agenda was instrumental, the District supportive, and Southview School active. The external aims and strategies were formative, while the internal aims and strategies focused on improvement.

The Nature of the Internal Accountability System at Southview School

The nature of Southview School’s internal accountability is reported here in answer to research question 1. The nature of internal accountability system is reflected in how the school accounts for the effectiveness of its instructional program. In Ontario, the EQAO recommends that schools account for their programs by consulting with school stakeholders, using assessment to determine student instructional needs, identifying school and community factors related to achievement, communicating about assessment purposes and expectations, and discussing achievement and improvement strategies. The study findings (condensed in Figure H9) on Southview School response to these five recommendations is described next.

Southview School was actively involved in *consultations with district and school stakeholders* to both establish and review school action plans. This process began with the School review audit when Mrs. Long spoke with parents, teachers and community leaders. She continued the consultations when reporting on EQAO results during the annual program review process aligned with the SIP. Close collaboration with district staff and other principals was also reported. Cooperative development of the details of the SIP by the administrative team and the teachers ensured that all staff members had input, thereby owning the change process. These actions indicated a strong response to the first recommendation.

The second recommended activity, *using assessment strategies to find out where students are in their learning and what they need to do next*, was established practice at Southview School. The Principal and teachers reported gathering and examining information on school performance from many sources. The EQAO test scores and school tests were used, but qualitative input from parents and community leaders was also taken into consideration in establishing goals. At the classroom level, an increased interest in using *formative assessment* was being fostered by new instructional materials in mathematics.

Identifying *school and community factors* related to student achievement was an inherent part of the school - community consultations that began with the School review audit and continued with the annual data gathering for program review purposes. Mrs. Long and the administrative team considered district socio-economic data, parental expectations, student grades and feedback from community stakeholders in interpreting school performance. These factors informed her expectations as the instructional leader at Southview School.

Teachers at Southview School *communicated about the purposes of assessment and discussed expectations* with students and parents. Teachers (Grade 4/5 teacher, F02, Grade 5/6 teacher, A02; Grade 7 teacher, F02) reported using rubrics and exemplars to demonstrate expectations to students and parents. Although all teachers did not use formative assessment to the same degree, all teachers reported that they discussed criteria with their students.

The School report and SIP process created the opportunity for principal and her teaching staff to *discuss school achievement and improvement strategies* with parents. The reporting and consultations among the school, district and community stakeholders involved a level of accountability that required explanation and justification of the instructional program and the improvement plans. The consultations with the parents framed shared expectations and resulted in community ownership for improvement plans. The School report provided transparency of school performance (EQAO scores) but did so in a manner that clearly showed how the information was used to plan for school improvement.

Just as the principal at Southview School discussed school-wide expectations with the Parent Council, she expected teachers to do the same with parents. She reported that she encouraged her staff to focus on the student learning and have ongoing goals that are shared with parents. She reminded teachers to inform parents that “this is where your child is at, this is where they need to go” (Principal, Ju02). However, teachers reported that during interviews and ongoing communication with parents, no one asked mentioned EQAO results, but rather discussed grade level expectations and what parent support was needed in terms of extra practice and home review (Grade 4/5 teacher, M02). The frequency of communication between home and school varied with the

individual teacher and the grade. Test scores and rubric assessments were the tools these teachers used to keep parents informed of student progress.

Together, the presence of these school strategies indicates a strong response by Southview School to the five recommendations associated with account making. In terms of transparent accountability, the first EQAO aim, Southview School was demonstrating a significant degree of implementation. In re-visiting the aims identified by the school in its annual report, “to provide an education of the highest quality,” Southview School was acting in ways that supported this goal.

In Summary

The strength of the internal accountability system at Southview School has been described as strong based on the presence of all five components of the integrated model. It supported school improvement through encouraging ownership for school improvement. Leadership to mobilize school staff and support the internal system was important. The depth and breadth of involvement in school accountability a process (review, SIP) was expanding, while at the same time was being managed by the principal to make small incremental but sustainable organizational change. School leadership was instrumental to the effectiveness of the internal accountability system. The EQAO testing was a second important influence on the internal system.

The compatibility between the aims and strategies at each system level (provincial, district, school) illustrated differing roles and responsibilities in the restructuring system. These differences reflected a balance between oversight (via substantive contributions) and local discretion (responsible decision-making). Furthermore, the compatibility was heightened by the fact that the priorities at

Southview (student achievement) aligned well with the external policy aims. The next section reports on research question #2 that sought to explore the influence of the extant internal accountability system to promote change in teacher practice, the second aim of Ontario's EQAO accountability agenda.

Influence of Southview School's

Internal Accountability System on Assessment Practice

This section describes classroom assessment practice and examines the influence of Southview School's internal accountability systems on teachers' assessment practices. Classroom assessment in mathematics is characterized in terms of stated purposes, criteria for selection, methods, and feedback. Then collegial dynamics that address the influence of internal accountability to promote standards-based assessment practices are described.

Classroom Assessment

Teachers' assessment practices in mathematics are described using four features including the purpose, the criteria, the method, and the feedback provided. These descriptions shape understanding of how teachers assessed and graded their students at one point in time during the EQAO implementation of the standards-based testing. *Figure H10* presents the findings in an abbreviated format and supports the analysis presented here.

Purpose

Teachers at Southview School assessed their students with different purposes in mind. These purposes were a function of when the assessment was conducted, why it

was conducted, and by whom. In other words, each purpose was determined by the *what* and the *why* an assessment was done. Southview teachers collectively reported six different purposes: diagnostic, formative, summative, student management, test-taking preparation, and accountability. Each is discussed in turn next.

Teachers reported that *diagnostic assessments* used both informal strategies and teacher constructed quizzes to learn what students already knew about the mathematics topics. The Grade 4/5 teacher (F02) began each mathematics unit by giving a short quiz to discover what students knew and to compare math knowledge between her grade 4 and grade 5 students. Often the younger students tended to be stronger academically and so her lesson planning took this aspect into account.

This pre-test really gives me a good idea, in general, where they are and what I should begin with in the unit. I do this because in some units this year, I've had to go right back to the basics. I had thought that they would have already learned the concepts and would be ready to move onto the new material in the curriculum. (Grade 4/5 teacher, F02)

The Grade 7 teacher began his math units with a class discussion and made note of students' comfort level speaking about the mathematics concepts. He found that "those with eye contact and those without, looking down and avoiding the teacher's gaze" gave him a good idea of student confidence with the skills and the terminology (Grade 7 teacher, F02). Diagnostic assessment in the junior and intermediate classes was primarily used to determine if review was needed or if students needed to be grouped for instruction.

The use of *formative assessment* was common practice at Southview School. It was primarily characterized by daily monitoring of assignments and small quizzes. The

Grade 4/5 teacher used formative assessment to focus students' attention on what they needed to learn.

Throughout the unit I'm always checking their homework or their class work that they do in their booklets and notebooks. I check that constantly. So if they are having trouble with something, usually, I'm pretty quick to catch it. So I find that helpful. I plan to have a couple of quizzes in each unit before the big test at the end. And, in this geometry unit for instance, where we are making solids and other constructions, I would also monitor the class to see how they were doing with the process of building the shapes. I found that showed me how quite a few students would look at the net, cut it out, and yet not be quite sure how to put it together. It was an interesting type of thing to observe. So this gave me an indication of what they needed to learn. (Grade 4/5 teacher, F02)

Formative assessment also involved students presenting their answers and checking their work as they completed it each day (Grade 5/6 teacher, A02). This teacher often selected activities from the textbook and its extra practice worksheets because these could be easily corrected. The new textbooks did raise awareness and foster change in the frequency of formative assessment.

Teachers at Southview School conducted ongoing *summative assessments* for each mathematics unit. The Grade 7 teacher (F02) reported that, "the assessment varies depending on the strand we are covering, but it is definitely throughout the entire unit. The ideal is to try and evaluate throughout using different methods. It is hard to do that sometimes but that is my intention." The final grade tended to involve a decision based on a review of all assessments conducted during the instructional unit including anecdotal observations. In some cases, a special unit end project was used as a cumulative assessment and would be given more weight in evaluating the student's achievement. The most recent, consistent level of performance recorded in a mathematics strand was identified and a judgment of overall achievement made. This way, the grade was "not based on just a final test" (Grade 4/5 teacher, F02).

I look at the progression and I find the level that they have progressed to. This is what their final mark is based on. (Grade 4/5 teacher, F02)

Curriculum achievement charts, rubrics and assessment guidelines focussed evaluation upon standard criteria, however, in assigning a grade teachers reported that subjective evaluations of student effort and commitment did play a role in their final decision.

I think that I am a teacher who errs on the side of higher marks, especially with Grade 5 and 6 [students]. If I'm deciding between a B+ and an A-, well you know, I will give him an A-. What's it going to cost? Encourage them. (Grade 5/6 teacher, A02)

There is a little bit of judgement involved there. So yes, I do consider the student if there is a 2+ or a 3-. There is some judgement there and it is very subjective. Do I consider the students? Like are they trying hard? I agree that effort would nudge up the grade a bit. (Grade 7 teacher, F02)

This suggests that the process of learning as captured in formative assessments was recognized and taken into account when an evaluation of summative assessments led to a decision about a student's overall achievement. Teachers acknowledged this subjective element in evaluation but explained that determining the grade was a conscientious process tempered by the provincial expectations and standards. The number and variety of assessments varied from teacher to teacher and the emphasis given to unit end tests depended on the particular math strand taught and teacher knowledge of performance assessment. Teachers reported that they followed the District assessment and grading handbook when determining marks for the report card.

Southview teachers found the report card format aligned with the curriculum expectations. The Grade 4/5 teacher (M02) reported that, "Our report cards must be done a certain way, and done well ... I find the report card ties into the expectations quite well. I can pull right from the expectations and put them onto the report and I know it

matches the levels.” Teachers said that the software program facilitated reporting by providing databases of expectation statements that could be personalized with specific phrases such as *a few minor errors* or *usually accurate*. They were able to add one or two lines that “say what I really think I need to tell the person or what I think he should do” (Grade 5/6 teacher, A02). This same teacher found that the technology was workable but the transition to the new format stressful and wondered if all the detail (5 separate math strands) was necessary.

Teachers (Grade 7 teacher, F02, Grade 5/6 teacher, A02) reported that planning instructional units to address reporting requirements had been problematic at first because the linear, sequential nature of the available textbooks didn’t align well with reporting requirements. They found it necessary to use multiple resources to organize their instructional programs, plan summative assessments and report on each strand twice within the school year.

Assessment at Southview School was also used to *manage* student behaviour. This purpose was evident in the transparent assessment process where rubrics clearly outlined the concept knowledge expected, skills required, and qualities of the final product.

I think there is an increase of self-confidence because of their achievement in the regular everyday work; if they see that they are being assessed on that as well as on the final test, or on a consultation. I find that when they do it well, their self-confidence increases. (Grade 4/5 teacher, M 02)

Teacher feedback focussed on next steps and students were recognized for the efforts they made to improve. Tangible recognition such as stickers on assignments rewarded the efforts of younger students. Teachers also recommended students for the monthly Math Achievement certificates and reported that selection was based on improvement

and hard work as well as achievement. For intermediate students, recognition sometimes involved recognizing quality products as an example to others.

Teachers reported that they assessed students using performance tasks from previous EQAO tests as they prepared for Grade 6 tests. This use began as *test-taking preparation* but also resulted in the use of the EQAO questions as part of the Grade 6 instructional program (Grade 5/6 teacher, A02). The tasks were found to be appropriate tools for determining student achievement against the curriculum expectations.

A final purpose for assessing students at Southview School was *accountability*. At both the junior and intermediate level, Southview students were shown rubrics and taught assessment criteria. This was done early in the school year because this strategy demystified the assessment process for the students and better informed their efforts. All teachers (Grade 4/5 teacher, J02; Grade 5/6 teacher, A02; Grade 7 teacher, F02) reported that they wanted students and parents to understand what was expected and what they should be working toward.

The Grade 7 teacher, for example, encouraged student ownership for learning by using a journaling strategy as unit review. It was a means to foster student reflection and accountability for their own achievement.

It is a way for the students to be a little more independent and go over the things we have discussed. It often seemed that the students used to complete their homework and move on to the next page and repeat this process [like] that until the unit was over. Then they wrote a test and it's finished. This way, they have to reflect on their learning, and if they don't, they have to look at their work and look at the textbook, read what's there and then summarize it in their own words. I hope that is more effective. (Grade 7 teacher, F02)

This teacher also used peers to promote accountability. He had students "evaluating the group presentation of a concept" (Grade 7 teacher, F02; M02), but what he described was a context of peer support rather than peer evaluation.

Criteria for Selection

Southview School teachers used various criteria when selecting assessment strategies. These are organized as five categories including relevance in terms of outcome, fairness in terms of comprehensiveness and objectivity, origin, relevance in terms of purpose, and their economy of use.

The teachers reported choosing assessment strategies, particularly unit end tests that *reflected the outcome taught*. They explained starting with the curriculum document and selecting instructional materials to address the outcomes listed. Some teachers (Grade 4/5 teacher, J02; Grade 7 teacher F02) selected performance tasks because they measured concepts, skills, problem-solving ability and mathematical reasoning collectively. This reflected the instruction and materials used during the unit. Journaling was a popular strategy (Principal, Ju02; Resource teacher, A02) and teachers used rubrics to assess use of mathematical language and notation to explain and justify reasoning. In other words, journal writing was an ongoing instructional strategy and assessment tool. One teacher continued to rely on traditional paper and pencil tests where the students “show their work and if they don’t, well then they don’t get as good a mark. So it’s [reasoning and communication] wrapped up in the mark. Then I take the mark and I give an A, B or C level. I can report it with confidence that the kid has gotten a string of Bs in his geometry strand” (Grade 5/6 teacher, A02).

A range of criteria for selecting assessments were reported by Southview teachers. Many teacher choices demonstrated pre-planning for formative and summative assessment to measure process and product outcomes; while one teacher seemed to focus on the end product alone.

Teachers reported that the curriculum documents, exemplars and textbook resources were good sources of assessments and this supported *fair assessment*. The Grade 4/5 teacher used the Quest 2000 program but did appreciate that an element of subjectivity was inherent in the holistic approach to grading in the assessment materials. She felt the District handbook guided how she evaluated student performance. The Grade 7 teacher (F02) reported some difficulty following the handbook guidelines when a student's performance in a unit was inconsistent since averaging the scores was unacceptable. He developed his own assessments when necessary to ensure that the questions matched the curriculum expectations closely.

Teachers at Southview School reported selecting *assessment strategies for different purposes*. These choices varied with the unit topic, geometry being cited as one math strand where summative assessments were often performance tasks designed to reveal process as well as produce a product (Grade 4/5 teacher, J02; Grade 5/6 teacher, A 02; Grade 7 teacher, F02). The Grade 7 teacher (F02) reported using math journaling for summative assessment purposes because he wanted to judge students' skill in organizing and explaining the math concepts and procedures. The Grade 4/5 teacher (J02), in contrast, explained that she used journaling and performance tasks as formative assessments because the feedback informed her planning and motivated her students to learn.

Another reason given for selecting particular assessments at Southview School was their *origin*. Document analyses of the sample assessments showed most instruments used by junior teachers (Grade 4/5 teacher, J02; Grade 5/6 teacher, A02) were taken from the new textbook and its resource packages. Some were pulled from the Ontario Curriculum Planner internet links and shared with colleagues. The Grade 7

teacher reported that he usually prepared his tests using a variety of sources, just as he had to pull together his instructional units from several textbooks in order to cover the curriculum expectations. All teachers reported adding basic computation questions from teacher math drill resources to tests to ensure basic skills were assessed.

Economy of use was a fifth reason for teacher selection. The junior teachers (Grade 4/5 teacher, J02; Grade 5/6 teacher, A02) reported using the assessments embedded in the math program because they were already prepared, used performance tasks and came with teacher friendly rubrics and checklists. These resources used a variety of strategies and the principal encouraged this diversity. At the intermediate level, the Grade 7 math teacher (F02) developed tests using several resources because no one math program met all the curricular expectations. This same teacher reported that finding appropriate assessment instruments was time consuming and he would welcome a more comprehensive textbook program.

Although not all teachers selected assessments for the same reasons, the practice of junior teachers was to select assessments and tests from the mathematics textbook and teacher resources because they aligned with the expectations of the curriculum and resembled the questions on the EQAO tests. This option was not available to the Grade 7 teacher who hoped that purchase of a new math textbook would be possible in the near future. Convenience and alignment with curricular outcomes seemed the uppermost criteria in teachers' choices.

Assessment Methods

Southview teachers primarily used three assessment methods to gather information about student achievement including question and answer, text-embedded questions and tests, and performance assessments.

Question-answer exchanges were consistent practice among Southview teachers.

The strategy was frequently informal and part of the daily lesson. The intermediate teacher kept anecdotal records of how well students expressed themselves in math class, while the Grade 4/5 teacher used question and answer conversation as a means to scaffold learning during performance tasks.

I spoke to every student and first asked them *What solid are you making?* to make sure they knew the name of it, then if they did I said, *Oh, that's great you knew the name. Excellent.* And *How many vertices does it have? Do you remember what a vertex is?* and if they did then, again I continued with that sort of back and forth dialogue....If they didn't know it well, I said *Well what else have we done here? Do you know what the edges are?* and if they said yes, that's great, I would continue on. So it was assessment and instruction at the same time. (Grade 4/5 teacher, M02)

Teachers (Grade 4/5 teacher, J02; Grade 7 teacher, F02) also reported using group interviews and student presentations to assess student performance. Classroom question and answer was usually informal and formative in nature.

Southview teachers used the *text embedded assessment* questions and accompanying paper and pencil tests from the teachers' resource package. These resources were accompanied by tools to help with the evaluation (checklists, rubrics, performance task protocols). These matched the scope and sequence of the textbook and showed a good alignment with the curricular expectations. However, junior teachers did modify assessments as the Grade 5/6 teacher (A02) explained: "Usually the culminating test is a mixture of both. So it will be the Quest 2000 final test and then I'll add in some of my own questions as extra pages or I'll take one of their pages and put in one of my mine."

The teachers also reported planning *performance assessments* to take into account the ability to apply concepts and procedures to arrive at a final product. Most

performance tasks came from textbooks, the EQAO test or the exemplar documents (Grade 4/5 teacher, J02; Grade 5/6, A02). Performance assessments were felt to provide a comprehensive picture of student achievement because “there are many students out there who don’t perform well on a test when they are actually quite comfortable with the work. Some students get anxious over tests and I think with projects, you get an overall look on how well or how poorly they know that subject area” (Grade 4/5 teacher, M02). Examining students’ ability to apply knowledge required “a formal evaluation structure” according to the Grade 7 teacher, especially when grading students holistically using the achievement levels 1, 2, 3, and 4.

Although performance assessments were common practice at the school, teachers continued to use paper and pencil unit end tests. This reflected an added-on approach to assessment. The preferred assessment methods used at Southview School included performance tasks, a variety of paper-and-pencil tasks (math journaling), and some form of oral questioning.

Feedback

Teachers reported that they gave feedback to students in several ways. The Grade 4/5 teacher (J02) preferred to verbally *praise* students for their effort to encourage them she gave stickers and wrote *very good* or *great* on assignments and quizzes. She explained that sometimes she “give them their grades and [but] the rest of the time it’s mainly positive feedback or else feedback of what I believe that they should work on at home” (Grade 4/5 teacher, J02). When the students worked on performance tasks, used praise to recognize effort then suggested next steps. She recognized that such *guidance* helped her students understand the steps of the learning process and effort was important. Her feedback appeared to be primarily *process-oriented*.

The Grade 5/6 teacher (A02) described a more traditional approach to feedback where class assignments were corrected and the scores spoke for themselves. On tests, additional marks were awarded for “showing their work,” revealing a preference for *numerical scores*. Math tests were marked as percentages and these scores were translated to letter grades in the report card. Formative feedback was therefore limited; the primary function of test scores was feedback final achievement.

The Grade 7 math teacher (F02) used several forms of feedback in his mathematics classes. Quizzes were marked with a *numerical score* and these were translated into a percentage. Praise and *descriptive comments* were written on assignments. The feedback tended to focus on the *product* except during conferences where “positive and sometimes negative” comments identified ways to improve learning.

Descriptive feedback supported student self-assessment and guided students’ learning because it used specific criteria to explain the learned skills and concepts while at the same time outlining what was needed to make improvement. Some teachers used descriptors for formative purposes in rubrics and checklists but it was common practice on summative assessments. On cumulative assessment tasks for example, feedback gave descriptions of performance in terms of effectiveness and was often accompanied by marks in symbolic, numerical or alpha-numeric forms. Praise in the form of words like *Great* and *This is a good journal entry* may accompany the mark. Both symbolic and descriptive feedback was reported in teachers’ assessment practices.

Classroom assessment practices reported by teachers at Southview School suggest a range of experience and understanding of the role of assessment to inform teacher instruction and support student learning. Assessment for grading was most

common at the grade 6 and 7 levels; while the Grade 4 teacher assessed for formative and summative purposes equally. Paper and pencil tests were common practice but all teachers reported using performance assessments and questions-answer exchanges as well. Teachers recognized student effort as well as achievement in deciding on final grades, but only the Grade 4/5 teacher verbalized the role of formative assessment and feedback in *scaffolding* student learning. Teachers were attempting to use a variety of methods and were familiarizing themselves and their students with evaluative criteria by using new assessment tools. The next sub-section describes the dynamics of assessment practice from a school perspective and identifies the professional interactions associated with how the internal accountability system influenced assessment practices.

Dynamics of School Assessment Practice

The second aim of the EQAO accountability agenda is to inform teacher practice. In order to determine the influence of the internal accountability system on teachers' practices at Southview School, teachers' assessment practices are characterized using five features indicative of collegial assessment: a) assessment goals, b) assessment instruments, c) assessment methods, d) personnel involvement, and e) values. A summary of the information is found in *Figure H11*.

Assessment Goals at Southview School

The assessment goals at Southview School reflected *compliance with the provincial standards-based assessment policy*. The principal and parents expected that all students should work towards the achievement of the provincial standard and that students be evaluated using the achievement chart. These *shared expectations* were

clarified at staff meetings as teachers participated in developing the SIP and collaborated on achievement targets.

The principal facilitated commitment to standards policy through leadership and through assessment conversations. Mrs. Long (Ju02) identified the provincial curriculum was a useful tool because it explained exactly what students needed to learn. All teachers interviewed reported using the curriculum expectations to plan their instructional program and confirmed using the achievement chart, provincially approved rubrics, and district assessment guidelines to evaluate their students' performance. Teachers felt the standard was challenging but achievable for most students at Southview School. School staff shared that special needs students (with IEPs) required additional supports to reach the standard and that parental support was needed to help with ongoing review. Concern about the achievability of the standard (level 3) for every student was expressed by the Grade 5/6 teacher (A02). It is important to note that this original resistance was assuaged as school performance on the EQAO and the individual achievement of his students consistently showed improvement. In fact, school performance over the three years the testing has been in place had consistently showed most Southview School students met or exceeded the standard.

Responsibility for teaching and assessing students to the new expectations required commitment by teachers. School performance on the EQAO, as an external measure, was not viewed as the sole responsibility of teachers in Grade 3 and 6 (Grade 4/5 teacher, M02). These *collective practices* indicate the acceptance of the provincial standards assessment policy at Southview School.

Assessment Instruments at Southview School

The internal accountability system at Southview School influenced teacher practice by locating, distributing and discussing assessment instruments and tools. The new mathematics programs possessed a variety of assessment methods and tools (rubrics, checklists) that aligned with achievement expectations. The EQAO anchor booklets, provincial exemplars and supplementary resources were gathered by school leaders and passed on to teachers at team meetings. These resources were sources of evaluative criteria and modeled differentiated levels of performance.

The use of *portfolios* was present at Southview School, but only as a tool for writing assessment. Mrs. Long shared that this was an area that she would definitely like to revisit next year. She knew that “all teachers have some kind of way to keep the records and samples of the children’s work. It’s not that the whole school does it the same way...but one of my goals would be student life conferencing with the whole school” (Ju02). Her goal was to introduce the use portfolios that held products of student learning and achievement across all curriculum areas to show student growth and development over time.

Assessment Methods at Southview School

Teachers at Southview School reported they were encouraged to use a *variety of assessment methods and tools* but were not monitored on how they developed and taught their instruction program (Grade 7/8 teacher, M02). As the principal, Mrs. Long explained, “classroom teachers have their different assessment tools and that can vary with some teachers that use math journals, some that do tests, like mini tests and unit tests. Some teachers do demonstrations, where math carries over into other subject

matter” (Ju02). She reported that teachers did not sit down as a group to develop assessments at this time, nor did she indicate that this was a potential expectation.

Teachers’ assessment practices were guided by the District assessment and grading *handbook* and the format of the standardized *report card*. Different assessment strategies were introduced at staff meetings and a variety of assessments were *embedded* in the new primary and junior mathematics program. New assessment methods were being *added on* to the traditional paper and pencil test.

At the time of the study, the teachers were immersed in learning the curricular expectations and interpreting them in terms of the new textbooks and the demands of the EQAO tests. *Teacher collaboration* among grade-level teachers at the junior and intermediate level was limited as each teacher sought to understand and implement the changes in terms of their own instructional program (English, French Immersion, split grades), leaving little time for informal collegial dialog during the school day. Collaboration was formalized when Mrs. Long shifted the purpose of staff and team meetings to address teacher learning needs and collegial planning. This provided an opportunity for a *common language of assessment* to be promoted and provided a forum to interpret and discuss the standard in terms of evaluative criteria. How this structured setting impacted on promoting teacher collaboration and assessment moderation could not be determined within the time frame of this study, but the internal accountability mechanisms were supportive of this type of collegial interaction.

Depth and Breadth of Personnel Involvement in Assessment

Discussions of school performance and instructional program among school stakeholders were reported at Southview School. These consultations, or rather

assessment conversations, began with the annual review of school performance and the development of the SIP.

The principal and her teachers met both collectively and in small working teams to evaluate the instructional program. The administrative team also worked with district staff, while the principal consulted with parents and community members to gather input and share information. All these conversations focussed on accountability and led to the shared expectations reflected in the school aims and the goals of the SIP. These assessment conversations were provided formative feedback because they guided the evaluation of the instructional program and the SIP process.

At the professional level, teaching teams selected and enacted the SIP activities, and school leaders promoted best practices. Collegial dialogue encouraged teacher experimentation with new assessments and encouraged use of rubrics and checklist to inform students about expectations. Parents played a supportive role with homework and review, but the development of individualized student learning plans was restricted to students with IEPs. Although responsibility for student grading was delegated to teachers and the breadth of involvement in SIP development expanded, the depth of staff involvement was limited. Teachers (Grade 5/6, A02; Grade 4/5, M02; Grade 7 teacher, M02) reported being stressed with the demands of implementing the new curriculum and having little time for reflection. As the Grade 7 teacher commented, “there is not a lot of time for reflection, frankly. There is time for planning and assessment, but no reflection.”

We certainly want to get as much consistency in marking as we can throughout all grade levels, that’s the first thing. And I really think that the main thing for teachers is to develop an awareness what each grade level expectation should be. (Resource teacher, A02)

At the school and community level assessment conversations were happening but not at the team level due to many factors, including program differentiations (grade 5/6 teacher, A02). The following quotes clarify the situation.

There is not a lot of opportunity to get feedback on my assessment. I am not as confident as I would like to be and I would like to learn more. (F02)

We have not gotten anywhere close to discussing any sort of test or assignment, or anything related to standardization in the school. (Grade 7 teacher)

However, teachers were talking about the changes “in passing”. The limited opportunity for reflection and collegial sense-making was prevalent among both beginning and experienced teachers;

A lot of teachers didn't have to do it before and were not aware of how to [use new assessments] ... if I go to another teacher and ask them, How do you assess your students on their performance in math?, they'll often say, 'You know, I am just winging it.' Because they just didn't have to do it before. So that's a hindrance. I think that a lot of teachers are unsure if they are doing the right thing, if they are assessing the right way, (Grade 4/5 teacher, M02)

A degree of teacher “isolation” was reported by all teacher participants and although they reported that the curriculum, the exemplars and the anchor books articulated the expectations well, there was uncertainty about the assessment process. Mrs. Long's formal organization of team planning at staff meetings was intended to initiate collaboration to begin to address these concerns.

Values at Southview School

Southview School possessed many qualities of an assessment community. School and community consultations facilitated involvement and ownership for school improvement. Team teaching was dynamic at the primary level. Collegial planning was not common practice among the junior and intermediate level teachers, although some

teachers expressed an interest in partnering and Mrs Long was beginning to structure grade-level planning meetings.

Some teachers spoke to the value of assessments for motivating students and scaffolding learning. Formative assessment was reported as part of teachers' classroom practices but the degree to which it was formal or informal (not recorded for grading purposes) reflected the teachers' instructional approach. Most teachers reported using a variety of assessments to take student learning styles into account. Teacher assessment knowledge was not yet at a point where teachers were completely comfortable with assessment practices but did *value the assessment process*. Assessment using performance tasks appeared to be *added-on* to traditional paper and pencil tests in the senior grades, suggesting inclusive assessment methods rather than selective decisions.

Formative assessment (Grade 4/5 teacher, F02) was used to individualize student learning through grouping and learning partners. It employed in-class arrangements rather than involving parents in the development of individual learning plans based on EQAO scores. Individual learning plans were designed for all students with IEPs and these plans were supported by the resource teacher and her team.

The presence of features characteristic to collegial assessment practices were reported at Southview School. The goals articulated by the principal and her staff suggest the qualities associated with an assessment community were desirable (Hall & Harding, 2002) and the principal reported that she has implemented organizational changes and strategies to support implementation of collegial dialogue and cooperation. Her activities relied on the mechanisms of the internal accountability system. Based on the events and procedures described collaborative assessment practice was not

established, but a transition from individualistic assessment to collegial assessment was in process and was supported by the internal accountability system.

Summary

The findings of this case study of Southview School support the proposition that a strong internal accountability system can be established in a nested accountability system designed to respect teacher professionalism. The aims and strategies of Southview School's internal accountability system reflected coherence with those of nested external accountability framework. Firstly, a strong internal accountability system is operating at Southview School, and school response indicates that Southview Schools is accountable to school and community stakeholders. Secondly, the internal accountability system influences change in teachers' practices. The characteristics of teachers' practices reflect the presence of formative assessment practices and acceptance of a standards-based approach. School assessment practice reflected many qualities associated with an assessment community.

CHAPTER 4

CASE OF NORTHSIDE SCHOOL

Northside School is a kindergarten to grade 8 elementary school situated within the same district and provincial context as Southview School. It is located within an urban community that possesses a socio-economically and culturally diverse student population. This case study of Northside School's internal accountability system explores the composition and strength of the internal accountability. It also examines its effectiveness to respond to the aims of the EQAO accountability agenda. The chapter is organized into two main sections that parallel the research questions. The first section examines the characteristics of the internal accountability system at Northside School, determines its strength, and then describes the relationship between the internal and external aims and strategies. The second section outlines the characteristics of assessment and describes the mediating influence of this internal accountability on teachers' assessment practices.

The principal at Northside School was Mr. Page. He was an experienced administrator who had worked at Northside School for several years. He welcomed participation in this study and assisted with the selection of the participating teachers. He and his staff presented themselves as teaching professionals interested in reflecting on their own assessment practices. The teachers at Northside School are identified by their teaching responsibilities in order to appreciate their contributions related to the role of the Grade 6 EQAO mathematics test on school accountability. The participant teachers included the Grade 7 mathematics teacher, the Grade 6 teacher, the Grade 4 teacher, and the Resource teacher who was in charge of the special education resource unit.

The Internal Accountability System at Northside School

This section of the chapter answers the three sub-ordinate parts of the first research question. This section first describes the component parts of the extant internal accountability system and determines its strength. Second, factors that influenced the internal system are examined. The third part of this section traces the relationship between the internal and external systems, concluding with observations about the nature of Northside School's response to system accountability initiatives.

Strength of Northside School's Internal Accountability System

This section reports on question 1a of the study. The characteristics of Northside School's internal accountability system are organized by the five component parts defined in the conceptual model and include: aims, information on school performance, standard of achievement, persons who judge school performance, and mechanisms to guide and monitor school practice. The findings are presented in a concise format in Appendix H (see Figure H12 for full proofs).

Aims

The first component of an accountability system, the school aims, indicate what the principal and its teachers are most accountable for to the public. The aims represent school priorities and reflect community expectations. As a *Signal School*, Northside School was socio-culturally and economically diverse with a student population possessing a high risk of failure. These social realities impacted students and influenced the instructional program. Mr. Page (Au02) described "a school that has every layer imaginable but with a concentration at the lower end. So probably a good third of our students, if not more, are living below the poverty line." The student population was also

multi-cultural and linguistically diverse since many students were landed immigrants. These special needs were supported by a special education support unit, several remedial support teachers, and a gifted enrichment centre. The principal reported that such complex needs required a coordinated and committed effort by all members of the school community.

Mr. Page explained that during his first year as principal of Northside School, he had begun an extensive consultation with all members of the school community to create a sense of identity and common purpose. He had initiated a vision building exercise designed to develop rapport and to define common values and goals by speaking with students, parents and teaching staff separately at first, and then sharing what each group contributed to the process with the others. Once all three groups had outlined important elements, the three versions were “brought to staff” and out of this, the final version of the school mission statement was generated.

The mission statement included phrases emphasizing *a safe, challenging and caring environment* and a focus on *human development*. The aims of the school were further explained in the School Report newsletter as: a) respect for the feelings, opinions and contributions of others, b) working toward one’s personal best, c) promoting high standards, d) promoting harmony and understanding, and e) educating for a well rounded life (School Report, 2001-2002). These aims are what the school was accountable for to its community.

Initially there were three issues of concern to the Northside School community, but “the first one was a school climate issue and we developed a school action plan to look at school climate, develop a common mission statement, and from there other products came” (Principal, Au02). At Northside School, the first focus was a

comprehensive plan to develop an inclusive culture and this led to the development of an extensive extra-curricular program involving student council, choir, homework club, breakfast club and a variety of special interest and sports activities. Safe school initiatives included school safety patrols, a social skills program, and a disciplinary framework based on maintaining student dignity. Teachers at Northside supported these programs and helped to plan assemblies and to distribute awards that recognized student efforts and achievements. The progress on school climate was reflected in the pride Mr. Page displayed when he spoke of the calm environment of the school, the staff interest in student well-being, and the culture of caring and respect. The positive impact of this aim was observed during school visitations.

After the development of the school mission statement, the School review process lead to a focus on academic issues, in particular student literacy. With English as a second language for many students, literacy was seen as a fundamental first step to student success in school. Encouraging students to do their best and promoting high standards in work and play were important aims, and these were the areas that Mr. Page and his teachers were accountable for in their instructional program.

Information on School Performance

The second component of an internal accountability system is the presence of information on school performance. The staff at Northside School collected a range of information to understand and judge the effectiveness of the school's instructional program. The data collection included the provincial EQAO test results and district information from the EDI, CCAT scores, and the Grade 7 math diagnostic test. To this information, statistical data on socio-economic and cultural factors that impacted the student population was added. This collection was consistent with the type of

information provided to other schools in the Central School District. However, the principal and teachers at Northside School sought additional information about community health issues and decided that additional diagnostic testing was necessary to address the unique needs of their student population. This represented a difference from other district schools.

Northside School decided to use the CAT as a diagnostic assessment each September in all classrooms. Mr. Page explained the reason for this decision.

The testing that we are doing, we hope will be diagnostic because we sometimes have kids for a short period of time. We felt that using our own testing would provide us with a way, very early on in the school year, [for] discerning which students need more individual attention and perhaps further screening. It also, at the same time, allows us to measure growth and in terms of scores, the teachers are using both grade scores and percentile scores. The percentile scores are being used to get a big picture and to be able to communicate to parents and to staff, take a look at that level in terms of data. The grade scores are going to be used to measure growth. (Principal, Au02)

The teachers supported this diagnostic testing because the information was timely and would help them form instructional groupings and plan instructional units. The intermediate teachers had been in the practice of using standardized tests in September for a few years already, previously using the language and computation subtests from the Gates-McGinty batteries (Grade 7 teacher, Ma02). They opted to use the CAT to align with the rest of the school.

The CAT testing was a screening tool to expedite the identification of students who would need remedial support to meet grade level expectations or who would require an IPRC determination to receive modified programming. With ESL and special needs heavily represented in Northside's student population, the Special Education team had seven special education teachers and five educational assistants. This team supported the

classroom programs or, alternately, operated companion remedial programs. The Resource teacher outlined the type of information that she gathered on students.

When I am assessing students, I am looking for areas of learning disabilities. I'm trying to identify learning disabilities, strengths, learning styles, so my kind of assessment is not content related. I think classroom teachers are doing the assessment for content materials, and knowledge and skills are part of content....The batteries I'll use [are] the WYATT, the Detroit Test of Learning and Aptitude, Peabody Language Development, the Nedemeyer Language Test. There are more individual tests. Some are ability, most are achievement, but they are standardized tests. (Resource teacher, A 02)

Together these sources of information provided an extensive collection of information on school and non-school factors that influence students' achievement and school performance. This information was used to plan instructional programs and to interpret school performance on the EQAO tests in light of all the information collected on students' aptitudes and achievement. This collection guided school improvement planning.

Standard of Achievement

The third component of internal accountability is a standard to measure achievement. At Northside School, the teachers and administrators respected the detail of the new curriculum, were familiar with the provincial achievement chart, and reported that they were in the process of clarifying performance criteria for the level 3 provincial standard.

Northside School teachers welcomed the clarity of the curriculum documents and the specificity of the expectations. They also reported that the EQAO tests and anchor books were helpful because these resources possessed examples of performance tasks and samples of students' work for each of the four levels of achievement. The anchor booklets showed what expectations looked like and detailed the evaluative criteria used

to measure a performance. Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02) revealed that these resources were, in fact, the main tool for learning about the new performance expectations and the level 3 provincial standard. With the distribution of the mathematics exemplars from the Ministry, these documents became another “favourite reference” for rubrics and criteria descriptors and “were used on an ongoing basis” (Grade 4 teacher, J02).

The Grade 6 teacher reported that:

I have grown significantly in my understanding [of] exemplars and levels 1,2,3,4. I apply them but loosely in terms of my personal expectations, understanding that certain students are very strong and some are very weak. I look for my individual expectations for the students. They have helped me a great deal in understanding about what the curriculum is looking for and I have studied quite a few assessment models for the EQAO and I try to apply those as best I can. (Grade 6 teacher, J02)

In reflecting on his classroom situation, a regular English program class, the Grade 6 teacher revealed that his students generally performed at a level quite different from the provincial standard. He felt that he and other staff were trying to “get on the same page” and were trying to “keep our expectations similar to that book”. As a result, he felt that at this first step in the transition “it is quite individual in terms of what you feel is appropriate and [if] you are meeting expectations” (Grade 6 teacher, M02). He sensed that student exceptionalities were recognized more at Northside School and were taken into account. This point of view was supported by a comment from the Grade 4 teacher. In reviewing her students’ results on the Grade 3 EQAO tests and cross-checking them with the Grade 3 report cards, she had noticed that “there is a movement towards being more lenient, which is easy to do in a school such as Northside” (Grade 4 teacher, M02).

This understood, teachers reported that they were better informed about the provincial standard and were working to more consistently apply the provincial standard in their classroom assessments. They understood that this was the measure by which they were to judge student performance and it was the measure used to judge school performance. The junior teachers at Northside were active in the development of collective expectations through team planning of assessment tasks and dialogue about the evaluative criteria (Grade 4 teacher, J02).

At the Grade 7 level, the math teacher reported that the Ministry exemplars were wonderful for demonstrating performance criteria and a “great reinforcement that you are doing the right thing and that your instincts are right on.” She explained that she discussed “what would be an appropriate level 1, 2, 3 or 4” with her team partner on a regular basis (Grade 7 teacher, Ma02).

Exposure to exemplars and anchors helped to build collective expectations for students in general and to address issues of specificity in setting appropriate expectations, but these expectations were often modified for students with IEPs. The Resource teacher, in describing her responsibilities at the school, explained that students with special needs worked toward different grade level expectations, “a more modified standard, keeping in mind that these are kids that have special needs” (Resource teacher, A02) and their instructional program and performance criteria may differ from grade level expectations. Teachers had to take this into account during the assessment and grading process and report these differences on the report card.

Teachers and administrators at Northside School understood the situation and their responsibilities and, while respecting the value of a standardized level of achievement to assist with objective assessments of school performance, the annual

program review took the characteristics of the student population and the unique special needs of the IPRC'd students into account when reviewing school performance and assessing effectiveness. A degree of variation was extant in the application of the standard in assessment and grading, but teachers were attempting to address consistency. Collaboration among teachers was leading to collective expectations based on the provincial standard.

Persons Who Judge School Performance

The fourth component of an internal accountability system is a decision-making body responsible for judging school performance. At Northside School this responsibility was a shared one and involved extensive collaboration among both school and community stakeholders.

Early in Mr. Page's tenure as principal at Northside School, he formed a teacher committee to coordinate and communicate improvement planning. The group was called the School-Based Decision-Making Committee (SBDM). The SBDM Committee was charged with the responsibility of developing a three-year strategic plan based on a set of hypotheses and a philosophical perspective rooted in constructivist theory. Mr. Page explained that his administrative team "asked for volunteers from each division and from each program. So on the team there is a primary, junior, and intermediate person, as well as a special education person and an ESL person" (Principal, Au02).

The Committee was formed so the improvement plans would consider different perspectives and appreciate that needs may be different at the primary, junior and intermediate divisions. Mr. Page was responsible for the SIP at Northside School but mentored teacher ownership for the program review (assessment) and the development

of improvement strategies based on conversations about school performance and shared expectations.

The SBDM committee analyzed the EQAO data, broke it down into specific subjects and strands to look at the trends. Mr. Page facilitated the review processes, but the SBDM Committee was responsible for the development of the SIP. The Committee also worked with district staff from the quality assurance division and met with the superintendent. Mr. Page explained the superintendent's role:

When we developed the plan, she [the superintendent] is present. And when we have the one-day session with quality assurance staff, she is present. She joins in the discussion so she is there at the inception stage and she has a tendency of asking some very timely questions. She is also providing some interesting insights. She gave us some insights into our literacy plan last year and shared a different way of looking at things. She triggered a great deal of discussion. (Principal, Au02)

The Parent Council also participated in the review of school performance. Mr. Page reported that Council members did not wish to participate in the SIP workshop but were interested in examining the EQAO results and wanted to understand what the scores meant. He reported that they looked at the SIP in detail and were supportive of it.

They did question different specific items and were seeking additional information. They are definitely interested but they are not overly stressed as a group with the scores....I think the key thing for them is that we look at the data, plan for improvement, and provide a challenging program for students. (Principal, Au02)

As principal, Mr. Page make sure that the SBDM Committee had release time to examine all data and had access to all information on school performance collected. This included feedback from the Parent Council. He facilitated an inclusive perspective because of such consultations. Mr. Page believed that "people have seen that there is movement in the right direction in terms of student performance but they are very cognizant of the fact that many variables that will impact on the actual scores are beyond

their control” (Principal, Au02). Mr. Page also encouraged his staff to take responsibility for reviewing effectiveness of the SIP. When the SBDM Committee shared the current EQAO data with the staff and discussed the trends that they identified from their analysis, the Grade 4 teacher (M02) who had previously taught the Grade 6 class explained “we were discouraged if the results didn’t come in as we had hoped. But I can honestly say that I didn’t feel any kind of pressure from administration. It was just more of my own, placed on myself.”

Program assessment at Northside School was multi-faceted because of the depth and breadth of school and community involvement. From the assessment conversations a shared sense of ownership for improvement planning was implied. How SIPs and other accountability mechanisms guided and monitored teachers’ practices at Northside Elementary School are reported on next.

Mechanisms to Guide/Monitor Teachers Practice

The fifth component of an internal system is a function of the instruments and strategies directing teachers’ practices. Several accountability mechanisms were found at Northside School. Some were formal, external requirements, while others were more informal, school-based routines and procedures. However, both types of mechanisms guided what teachers did and revealed how decisions were made. These mechanisms are detailed here.

External mechanisms. The external mechanisms at Northside South included the School Report, the SIP guidelines, the provincial report card, the district assessment handbook, the TPA guidelines, the Ministry target-setting and the School Audit process.

One external mechanisms associated with the EQAO was the annual *School Report*. Like all Central District schools, the Northside School principal prepared and

published a *School Report* newsletter that was updated annually and sent home to parents. The report outlined information about the school and its aims. It described the community, the student population and the teaching staff. The instructional programs were listed and key school-based initiatives described. This newsletter, written by Mr. Page, outlined the “measures of student achievement and success” as well as the “school improvement plans and initiatives.”(*Northside School Report*, 2001-2002) The EQAO results were published as aggregated percentages of students achieving each of the four levels of achievement for each sub-test (reading, writing, mathematics). The following qualifying comment introduced the results section.

This assessment by the Educational Quality and Accountability Office (EQAO) is only one measure of the many ways in which teachers measure student performance, and teachers will be looking at the results in the context of other information available about each child. Results on this one assessment may not reflect a student’s performance during the entire year. The same is true for the school-level results reported below; they should be interpreted in the context of other information, some of which is available in this school profile.” [Northside School Report, 2001-2002]

All parents received a copy of this newsletter and it was published on the district’s website. Parents also received a report of their child’s individual scores, but classroom teachers reported fielding no questions from parents about individual student results. As the Grade 7 teacher (Ma02) explained, “I do not think scores can be the responsibility of only one teacher.” Mr. Page (Au02) agreed and explained that he didn’t feel that schools were “held accountable in terms of specific results or scores.” Responsibility for the instructional program was another matter and the newsletter was an opportunity to share information about school programs, stated aims, local priorities, and improvement plans.

Northside School was accountable for interpreting the EQAO results and other school performance information as part of the program review process. Performance information was discussed with the Parent Council, community leaders and the superintendent.

The second external accountability mechanism was the *School Improvement Plan*. At Northside School, the development of the SIP was ultimately Mr. Page's responsibility but he delegated much of the work and responsibility to his SBDM Committee and through them to the teachers. The Committee attended a one-day workshop where District staff explained the guidelines for SIPs and provided expertise on how to analysis the EQAO data and other relevant information. The Committee learned how to make comparisons, establish trends, review options and consider action plans that would best meet the needs of the Northside student population (Grade 4 teacher, M02; Grade 7 teacher, Ma02). SIPs from other schools were available to stimulate reflection. Mr. Page explained how the superintendent is involved.

She does take a look at it. It is presented so that there is discussion and she does follow up over the course of the year to see how things are going. Last year, we met twice to go over the plan and see how things were progressing with the implementation of the plan and she offers assistance. It is part of, but not the sole focus, but it is certainly something that is considered at each of those meetings. (Principal, Au02)

The teachers at Northside School reported the experience worthwhile and indicated that the staff was happy with the decisions the Committee made.

The *SIP guidelines* framed district objectives and initiatives, listed the stages of the planning process, required specific outcomes for literacy and numeracy to be identified, suggested resources, and provided a template that formatted activities with responsibility centre and deadline dates. These supports laid out a process, but as Mr.

Page confirmed that as far as the actual content of the plan “we need to focus on two areas, literacy, numeracy and / or a third. So there is direction but within the numeracy and within the literacy, we are not given targets or specific areas to focus on” (Principal, Au02). Northside School made school climate a third initiative. For the purposes of this study, the mathematics section of the SIP is examined further. Northside’s SIP listed the following numeracy outcome and related activities:

The % of students performing at or above level 3 will increase:

1. Numeracy SBDM Committee to develop action plan for 2001-2002 with emphasis on promoting understanding of concepts and communication of required knowledge (vocabulary development).
2. Gather and analyze student data to include: diagnostic assessments of all grade 1, 2, 4 – 8 students spring and fall; EQAO data; report card scores and demographics.
3. In-service staff: teaching learning strategies that promote vocabulary development, communication of ideas, and effective problem solving.
4. Inventory math support materials and acquire math manipulatives.

[SIP 2001-2002]

The SIP was reviewed annually when the EQAO results arrived. School performance was averaged over three years then compared to district and provincial aggregated scores. Northside School averages were consistently below district and provincial averages of students meeting or exceeding the provincial standard.

Mr. Page acknowledged that ongoing improvement was a school expectation and improved performance was anticipated. He cautioned that “one third of our students change over from year to year, we did set targets and we are looking for growth each year but we have been conservative” (Principal, Au02). The SBDM Committee drafted the SIP after conferring with the divisional teams about strategies and resources.

Setting modest but achievable targets was motivation to remain assertive and forward thinking even in a school context where so many variables affecting student achievement were outside the influence of the school. The principal indicated that goals

focussed staff attention on aiming for improvement but concentrated energies on those variables (instructional program) that they could influence.

The third external mechanism that guided and monitored teacher practice at Northside School was the *provincial report card*. Teachers (Grade 4 teacher, M02; Grade 6 teacher, M02) found that the report card format structured reporting on the five mathematical strands. The software program enabled teachers to use global statements to frame the curriculum expectations and then fine tune each comment to each student's level of achievement. There was a direct link to the standards (Grade 7 teacher, Ma02).

The *district handbook for assessment and grading* guided the reporting process by requiring each strand be graded twice each school year. This frequency meant the curriculum needed to be taught in a more comprehensive manner. Mr. Page reported that changes in mathematics instruction were happening.

I think its implementation in terms of making sure that every strand is looked at twice a year, is happening. So we are changing the focus from *They can multiply, add, divide and subtract* before we do anything [else]. That focus is finally moving away and teachers are spending the time that needs to be spent with the geometry and the measurement and the probability. (Principal, Au02)

The report card format and the district handbook held the teachers accountable for teaching the mathematics program as the curriculum intended.

A fourth external mechanism was the *teacher performance appraisal (TPA)*. The Ministry of Education had revamped the teacher performance appraisal and sent new directives to school districts in early 2002. Central District had a detailed performance appraisal system in operation that, like the new provincial TPA, was based on the College of Teachers' Standards of Practice. According to Mr. Page, the district's appraisal system was already aligned with the Standards of Practice and he was not expecting any changes in terms of the content of the appraisal.

After attending an administrative meeting on the new performance appraisal process, he found that “ the one component that’s drastically different, in terms of data gathering, is a standardized format of gathering data from parents” (Principal, Au02). The College was revisiting teacher recertification and appraisal processes and made a decision to provide parents with an opportunity to comment on teacher’s performance. Mr. Page was not yet clear on what type of input would be collected, how it would be collected, or what impact it would have on teacher evaluation. Although Mr. Page did not see the appraisal process changing much, it is fair to mention that several teachers had TPAs done in May and June just after the principals’ workshop but before the new guidelines were implemented. The Grade 7 teacher (Ju02) shared that her performance had been appraised, something that had not been done in some time.

A fifth external mechanism newly introduced during the period of time that this study was conducted was the *Ministry target setting* exercise. This exercise involved the principal and his primary teachers and it required that teachers assess each student from Kindergarten to Grade 2 to determine the current level of achievement in reading. The aim was to set a goal for student performance on the Grade 3 EQAO reading test. According to Mr. Page (Au02), the focus was not “on the actual scores” but on seeing an “increase in the number of students who are reaching the standards.” Each year individual student progress would be monitored. Program effectiveness would be assessed based on students’ actual performance on the Grade 3 EQAO reading test. Student grades and CAT test results would serve as annually measures of progress.

The sixth external mechanism reported here is the *School Audit* that was carried out when Mr. Page assumed the principalship at Northside School. It was a key district mechanism formalized when the Central District put its Strategic Plan in place in 2000.

Although Mr. Page did not refer to the audit processes specifically, the development of the school mission statement and the formation of the SBDM Committee reflect outcomes from it. Another outcome, the principal's strategic plan, is reported further in the next section on the internal mechanisms of Northside School's internal accountability system.

Internal mechanisms. Six internal accountability mechanisms linked to Northside School's organizational management are reported here. These include the strategic management plan, the school-based accountability testing, annual growth plans, new instructional materials, staff meetings, and long-range plans.

At the Northside School, the *school strategic plan* is one of the internal mechanisms that guided teachers' practices. As principal, Mr. Page had three management outcomes. He did not state these explicitly but they were implicit in the way that he organized his staff and school resources. The first, addressed above, was the development of the mission statement, the second was staff renewal, and the third was instructional change. The latter two goals were integrated. This was done by the formation of the SBDM Committee.

Mr. Page met with the SBDM Committee to develop hypotheses and to set parameters that reflected "certain realities within the building" (Au02). As principal, he provided a philosophical direction and brought resources and expertise to assist with the planning process. The SBDM committee worked with the following assumptions:

1. classroom improvement, teacher development and school improvement must be systematically linked (SIPs, in-service and supervision for growth are linked)
2. the *teacher as a learner* concept is the key to linking classroom and school improvement (personal growth plans)
3. small but persistent changes are more apt to endure (grounded in data and needs)

4. that frequent feedback loops allow for adjustments, recommitment and celebration (SBDM Committee Report, 2001)

These assumptions revealed professional ownership for school accountability. The role of the principal was to set the structure, identify the important guiding principles, provide input and expertise, and then facilitate the teachers as they analyzed needs, outlined strategies, and developed collegial venues to carry out the plans. Mr. Page (Au02) reported that he reviewed the plan when it was developed by the SBDM Committee and discussed the contents before it was shared with staff and parents. The multi-year plan worked with three external initiatives: Central District's School Audit for excellence in curriculum and program, the Curriculum Contact Teacher (CCT) initiative, and the SIP mandate.

The school strategic plan incorporated actions that were both organizational and instructional in intent. The former reflected management dynamics while the latter were leadership functions. The formation of the SBDM Committee characterizes a collaborative approach to school leadership, one in which teaching professionals played an integral role in guiding teachers' practices.

Organizational management was the principal's responsibility, but one that Mr. Page chose to manage through collaborative processes as much as possible. This is not to say that he didn't facilitate the strategic plan by making specific decisions related to staffing. He sought to integrate experienced staff with recent faculty graduates, having recently hired two experienced junior teachers, a highly qualified resource teacher and a first year teacher. This was done subtly as revealed when Mr. Page (Au02) discussed his management objectives within the *staff hiring process*.

When Mr. Page interviewed for vacant positions, he gathered information on teachers' knowledge, skills, and attitudes regarding curriculum planning and assessment practices. He reported looking for teachers who were aware of the changes happening in teaching and tried to hire people that had experience with new teaching strategies and assessment practice. He explained how this was done in the following comment.

In the interview, we talk about how does someone go about planning a unit. How does someone go about modifying the program to meet the needs of students who learn differently? We ask specifically about the area of assessment, *How do you assess and communicate the information that you gathered with students?* For the most part, we are having much better answers to the questions that we have been asking for the last few years. The way the questions are phrased, we are looking for behavioural answers so that people are able to draw on experience and cite specific examples of things that they have done from the past to illustrate how they are going to do these things. So from a hiring point-of-view, we haven't seen much renewal but we are seeing better answers. The new teachers who are coming into the school do seem to have a better handle on these areas and the staff that has been here for 20 or so years are more aware of the curriculum expectations, the process and the assessment. (Principal, Au02)

This strategy was supported by comments from the Grade 7 teacher and the Grade 4 teacher, who both reported that they were teaming with teachers relatively new to the teaching profession. The Grade 4 teacher (M02) found that her colleague brought new ideas about assessment from her professional certification program and found these helpful when planning instructional units.

A second internal mechanism that guided teacher practice was the *annual growth plan*. Mr. Page explained that he asked each teacher to target at least one of the SIP goals in their professional growth plan each year. He found that the most common was literacy and he speculated "perhaps it is because of the way we developed the plan, starting with the school climate, then we went to literacy, then we went to numeracy"

(Principal, Au02). In general, Mr. Page felt that there was fair amount of renewal in professional learning.

A third internal accountability mechanism was the Northside School's *CAT testing* initiative. One key outcome of the program review process (SIP) was the SBDM Committee suggestion that all teachers give the CAT literacy and numeracy sub-tests each September as a diagnostic tool. The Grade 4 teacher supported the decision to do the CAT tests and saw its value as a pre-assessment tool in September but was "not too sure, in all honesty, how the April one is going to be used. The reason I am fuzzy about it is because I was actually interested on the September one" (M02). Mr. Page was able to clarify the rationale from his perspective as principal and a member of the SBDM Committee.

We will use the grade scores pre and post, to see what type of impact we have had over six months. That was the intent. Will it provide us the data we need? We'll have to see in March when we do the actual retesting to see what the situation is. (Principal, Au02)

Mr. Page (Au02) elaborated that the SBDM Committee wanted to keep the testing reasonable so "we have given the grade 3 and 6 classes the option of not redoing the test in the spring because they are going to be doing the EQAO testing. But all grades have done it in the fall." He reported that testing gave information on student achievement in a timely manner and was used by teachers to develop their instructional programs. This strategy would also give feedback each year for each grade and would allow more pertinent feedback on assumptions made about the EQAO mathematics performance being a reading problem.

A fourth internal mechanism that guided teacher practice at Northside School was the purchase of *new instructional materials*. New mathematics textbooks were

purchased when the EQAO results showed that students' performance on some of the mathematical strands were very weak. The SBDM Committee determined that the teachers were using outdated textbooks and were not accustomed to using math manipulatives. Representatives from publishing companies made presentations at staff meetings and "Quest 2000 was chosen by the staff" (Grade 4 teacher, M02). The intermediate teachers chose a different series, one that was more balanced in its "use of language" (Grade 7 teacher, MA02).

The new textbooks used a language based problem-solving approach that some Northside teachers (Grade 6 teacher, M02; Grade 7 teacher, Ma02) found demanded strong reading and communication skills. The materials integrated several concepts and used performance tasks. This posed some difficulty in the beginning when teachers were learning to introduce topics and plan lessons on specific mathematical concepts.

The Grade 6 teacher (J02) explained that he found pacing himself to cover all the strands and using the new teaching materials a challenge. In his words, "I am trying to stick with things that I am comfortable with, confident with, and as I become more familiar with the book and the curriculum, I am going to start new assessment techniques." He persisted with using the new mathematics resources because it modeled the type of problem-solving activities found in the EQAO tests.

Staff meetings were a fifth internal mechanism at Northside School. These were venues for professional development because teachers were introduced to best practice and were asked to share instructional strategies and assessments. Staff meetings and team meetings were each held once a month. Mr. Page and his SBDM Committee used these meetings to communicate priorities and keep improvements plans on track. Examples of professional development reported by teachers included presentations by

textbook publishers, demonstrations of hands-on ideas for mathematics, and a presentation by the junior team leader on math journals and rubrics. This strategy was low cost, school-based professional development that did not require teachers to give up personal time after school hours.

Long-range plans were the sixth mechanism that both guided and monitored teacher practice at Northside School. Mr. Page explained that teachers' plans were used to monitor the instructional program, particularly as it related to special needs students working in modified programs.

In the long-range plans we do expect to see the type of assessment tools that will be used. We do ask that a range of assessment strategies be employed. We do have a range of skills within the staff in terms of who is using a more comprehensive approach. Some are more traditional with their assessment strategies. (Principal, Au02)

He explained that it was important to guide teachers on working with IEPs and planning for special needs students. The plans were one way to monitor if teachers used alternate methods to assess their students. He felt that this was one area that the school needed to work on. These plans provided a snapshot of teachers' practices.

In total, these guiding mechanisms are comprehensive and reflect the degree of internal accountability operating at Northside School. Some mechanisms required teachers to provide an account of what they did, while others required teachers to follow guidelines that directly influenced the design of their instructional program.

The internal accountability system at Northside School possessed all five components of the integrated model proposed. Employing criteria from the conceptual model, the internal accountability system at Northside School is strong.

Factors that Influence the Internal Accountability System at Northside School

This sub-section addresses question 1b of the study and describes the environmental factors that impacted Northside's internal accountability system. For the purposes of this study, the main factors included policy requirements and the mobilization of supportive actors. Two key elements (see Figure H13) were the EQAO tests and school leadership. These are described next.

EQAO Performance Tests

The EQAO tests influenced internal accountability at Northside School in several ways. They were an objective measure of school performance, models of performance tasks, and influential to what teachers did and why. Five specific impacts at Northside School are reported here.

Mr. Page and his staff recognized that the EQAO provided an *external measure* of school performance. This single test provided feedback on the instructional program and when scores were analyzed taking other information into consideration they were useful in school improvement planning.

The data we looked at initially was the EQAO and the evidence that we cited was that 42% of our students scored at or above the provincial standard (level 3) and 29% of our grade 3 students scored above level in problem solving and 28% of our students scored at or above level in the communication of required knowledge in math. So we targeted those areas as something to look at and try to improve our scores. (Principal, Au02)

These EQAO results were standardized and could help determine if the goals and activities in the SIP were improving the effectiveness of the instructional program. Mr. Page (Au02) revealed this impact when he explained that the school was "already seeing a trend in the data- that our students are not doing well in problem-solving" and this was a "surprise" because the scores in reading were improving. Mr. Page and the SBDM

Committee were operating on an original hypotheses that math problem-solving was closely linked to reading skill. The test results suggested that this assumption needed to be revisited because improvement in students' reading performance did not carry over to improved performance in mathematical problem-solving. The EQAO testing impacted on internal accountability in other ways as well.

A second impact of the EQAO tests was *curriculum renewal*; in mathematics this meant teaching all five strands. The Grade 6 teacher (M02) reported that he now addressed the expectations for all five mathematics strands because he knew that he hadn't been "hitting all the areas in my first two years so there was a lot of catch up before the test so that at least they were prepared." He sensed that curriculum renewal was happening at every grade level because he found that he could now "write a fraction on the board and they know what a fraction is" (Grade 6 teacher, M02). This hadn't been the case in previous years and he attributed this to changes triggered by the EQAO testing.

The Grade 4 teacher supported this influence and explained that she was much clearer about what she had to cover than in previous programs. She thought that class time for mathematics was adequate but required careful time management and program planning because of the comprehensive curriculum. As she explained:

I find differences in the way that I am going about things and I am certainly much more aware of the amount of time that I have to cover all of this because I find that the math program is a very intense one. So I find myself constantly looking at what do I have left to cover and how much time do I have to do that. Whereas with previous years, before the new curriculum, I did not feel nearly as a ware of time restraints related to how much I had to cover. (Grade 4 teacher, J02)

A third impact of the EQAO testing was an increase in *teacher ownership* for student achievement and reflective practice. This comment by the Grade 4 teacher captures this point.

I am always looking at- *What did I do last year or this year that didn't work out?* or *How can I improve upon it?* So if you ask me next year, I would probably have improved upon my program [*sic* but] would I answer the question any differently? No, I am not really satisfied. It is something about continually learning and growing from the point at which one is at. I don't feel that I am lacking for materials or information. (Grade 4 teacher, J02)

The testing also triggered collective ownership for student achievement and school-wide accountability for the instructional program. Having taught a grade 6 class the year before, the Grade 4 teacher compared her sense of responsibility between the different grade levels.

I still feel pressure down here in Grade 4. One of the difficulties of the EQAO [testing] is that you feel that you must really cover all the strands in order for the children to have a fair chance at the EQAO test. I felt more pressure last year in Grade 6 than I suspect I will and have felt this year in Grade 4, but there is no huge difference. (Grade 4 teacher, J02)

The message from the school leaders was clear- student performance on the EQAO tests was not just the responsibility of the Grade 3 and 6 teachers.

A fourth impact of the testing was experimenting with *new instructional resources* including assessment tools. The Grade 6 teacher (M02) reported that every grade level was trying to use the performance tasks from the new textbooks. To understand what students should be prepared to do on the tests, teachers were also using assessments from the EQAO tests, anchor books and exemplar documents. These modeled different assessment methods and articulated evaluative criteria.

A fifth impact of the EQAO tests that was particularly noticeable at Northside School was the increased attention given to the early *identification of special needs*

students. Northside possessed a mobile student population, many of whom required accommodations and program modification to meet their learning needs. Teachers (Grade 4 teacher, M02; Grade 6 teacher, M02; Grade 7 teacher, Ma02; Resource teacher, A02) reported that the EQAO was stressful for many students because all accommodations could not be adequately addressed during the testing period. The perceived impact on student esteem was a professional dilemma. As teachers found organizing the instructional program meant careful use of teaching time, finding time to give remedial help was limited. One solution was the comprehensive screening of students each September using the CAT reported earlier.

As a result, teachers at the Northside began to put more emphasis on identifying weak students early in the school year. The Resource teacher (A02) was instrumental in in-servicing the staff on how to modify classroom instruction, getting learning disabilities identified and organizing remedial support. She explained when to use the report card qualifier boxes that indicated modified programming (ESL, IEP) and applied for grant monies to support special programming.

Leadership

Leadership was an important influence on school-based accountability at Northside School for two reasons. The first reason is that Mr. Page valued teacher professionalism and was able to establish good rapport that facilitated the effectiveness of the internal accountability. The second reason is that he used collaborative processes to build commitment to change (professional learning).

The formation of the *SBDM Committee* was one way that Mr. Page's leadership mobilized support for internal accountability. It was comprised of administrators, program staff and teacher representatives from each division. This group sought advice

from district experts and consulted with the superintendent as part of the program review process. The SBDM Committee (Grade 7 teacher, A02; Grade 4 teacher, M02) also collaborated with the primary, junior and intermediate teachers during the development of the SIP. This enabled *shared expectations* to be built and *professional responsibilities* to be understood.

At Northside School, Mr. Page promoted *team teaching* and *mentoring* among the staff. Primary, junior and intermediate teams met regularly for professional learning activities and for collegial planning. The principal organized mentoring relationships when new teachers were hired (Grade 4 teacher, M02) or when teachers changed grades (Grade 7 teacher, A02). Some release time was provided and mutual preparation period arranged when possible. Northside teachers (Grade 7 teacher, A02; Grade 6 teacher, M02; Grade 4 teacher, M02) were positive about these partnerships and maintained that these helped cope with the stress by providing an opportunity to bounce ideas about, vent frustrations, arrange extra help for students, and maintain a sense of humour. The Grade 4 teacher (J02) found that “there is a lot of sharing, and there is quite a bit of teaming that actually happens at the junior level.” The same was reported at the intermediate level (Grade 7 teacher, Ma02).

Mr. Page recognized the value of *professional networks*. In particular, he was instrumental in establishing a Signal School Network, a group that met annually to discuss the unique issues affecting schools with high risk of failure populations. He encouraged his vice-principal to chair the network meeting and saw this venue as a means to exchange ideas and “hopefully to try and make connections with other schools that are dealing with the same type of issues” (Principal, Au02).

Mr. Page's leadership focussed on building collegial structures and mobilizing supportive actors to encourage experimentation. Mr. Page (Au02) supported the accountability initiatives, however, teacher morale was a concern because he found that "teachers haven't been given the time to sit down, digest, think and personalize the knowledge in this new curriculum." Teachers (Grade 4 teacher, M02; Grade 6 teacher, M02, Grade7 teacher, A02) appreciated the principal's support and as the principal had pointed out, Northside School did not experience a major turnover of staff.

The external EQAO tests and leadership at Northside School were two environmental factors that enabled the internal accountability system.

The Relationship with the External Accountability System

This section reports on research question 1c, the relationship between the internal and external accountability systems. For the purposes of this study of Northside School's internal accountability system, this section will address the relationship from the school perspective. The elements that comprise the components of the nested context are described here (see Figure H14 for full proofs). The detailed descriptions of both the provincial and district system components are presented in Chapter 3 (see pages 144-146) and are the same for Northside School. For this reason the external components are not reported here in the same depth but the degree of compatibility with the internal components at Northside School is addressed.

Compatibility of Aims of Accountability

Northside School's aims are reflected in the mission statement developed through a collaborative process with the students, parents and teachers. They are also reflected in the SBDM Committee's three-year strategic plan that frames the learning

outcomes addressed in the annual SIP. The aims focus on school climate and a positive learning environment where students are encouraged to respect diversity and to work towards one's personal best.

Mr. Page spoke of the calm, respectful environment of the school with pride. He also spoke to the literacy initiatives that he and his staff had put into place to address his large ESL population. These two aims were of highest priority in the first two years of the SBDM strategic plan and represented the areas where his school were most accountable to the parents and the community.

The specific literacy and numeracy outcomes listed on Northside's SIP (2001-2002) targeted increasing the percentage of students achieving level 3 on the EQAO Language and Mathematics tests. The district's aim to focus school improvement effort on student learning goals through its literacy and numeracy initiatives support the EQAO aims. This support was tangible in the SIP guidelines and the technical support for the School Reports.

Compatibility of Information on School Performance

All three levels of the nested accountability context supported the use of multiple sources of information to measure school performance. The EQAO provided an external measure of achievement, the Central School District provided other testing data and, in the case of Northside School, important socio-economic and cultural information to assist the school understand the diverse needs of its *Signal School* population. The District's initiative to identify schools with large at risk population was an outcome of System Review audit. To this extended collection of data, Northside School also added in school wide diagnostic testing using batteries from the CAT.

Compatibility of Standard of Achievement

Northside School was knowledgeable of the provincial standard of achievement and used it to inform their assessment and grading practice. Due to the number of students with IEPs at the school, school expectations for many students differed from this benchmark. However, any changes in how students were assessed or graded was defined in the IEP and reported on the school report card.

Compatibility of Persons Judging School Performance

The program review was a collaborative process at Northside School involving the SBDM Committee, the superintendent, and the principal as a participating member. Just as the collection of information on school performance was extensive, so was the depth and breadth of consultation about expectations, measures of improvement and identification of next steps. Parental input was also sought and welcomed as was information from community support agencies. This process engaged school and community stakeholders and was compatible with the intentions of the external system.

Compatibility of Accountability Instruments and Strategies

The external accountability instruments and strategies for Northside School were similar to those of other Central District schools except for one district level strategy that took its exceptional community characteristics into consideration. Several internal strategies operant at Northside School are also reported here because the degree of compatibility invites comment.

District instruments and strategies. Central District developed a special index that flagged schools where students exhibited high levels of poverty, mobility, cultural and linguistic diversity, readiness to learn, needs, and low socio-economic standing. The *Signal School identification* was used to address issues of equity and fairness in

education. The purpose of this identification was to improve opportunities for students and to foster success in schooling by seeking additional supports from the Ministry of Education and additional funding from civic, provincial and federal grant projects (Central District, October 2000). The Signal Schools initiative fostered dialogue between district support staff and teachers at the school. It facilitated partnerships between Northside staff and community agencies and cultural support groups so that a coordinated support structure was available for students and their families within the school community.

One tangible benefit was the assistance of language liaisons to help teachers and non-English speaking parents communicate about children's progress. Another benefit was district support for its large Special Education Resource Unit (17 persons) and a third benefit was a close working relationship with community support agencies like the Child and Adolescent Health Division of the city Health Department.

Northside School instruments and strategies. Northside School used many instruments and strategies to ensure accountability for its program and teachers' practice. This included additional school-based accountability testing (CAT) to support SIP initiatives and better accommodate student programming needs. Another issue related to compatibility of internal and external strategies was the participation of special needs students in the EQAO testing. These issues are unpacked here.

The CAT testing helped the school address the issue of student mobility. Mr. Page (Au02) reported that Signal Schools like Northside School knew that many students may not be in the same school for more than five or six months. Many students moved between schools in the Signal School family and principals were trying to find ways to help these students because, as the Resource teacher (A02) explained, they often

fell between the cracks when it came to screening and full identification for special education support.

Student mobility meant it was important to identify gaps in learning and special needs early in the school year. The CAT testing was one means to do this. The Grade 7 teacher (A02) reported that when she identified six new students in her math class with significant needs, the resource teacher was able to find another teacher with three “down time” periods who would run a remedial program for the group.

Although the benefits of the EQAO results to inform program review were understood and appreciated by the Northside School staff, dilemmas did arise with the nature of the EQAO tests themselves because current policy required that all students in Grade 3 and Grade 6 write the tests. The requirement was problematic because IEP'd students who worked with modified expectations or accommodations took the same tests as their classmates.

The conundrum was that accommodations (scripting, extended time, quiet environment, and reading support) for so many students were not always possible to provide. No accommodations were made for reading level or terminology, and the Grade 6 EQAO tests were written at the grade 6 reading level. As the Resource teacher (A02) explained, “this particularly affects the math results because we may have a student who is able to do the math but cannot read the math work, the math questions on the EQAO test.” She stated her position as follows:

We have a need to exempt our special needs students. They won't be, they can't be successful in the format that it is presented. They are used to working with support and we have taught them to seek help, to be an advocate for themselves, to request help when they see that they need it... They can't use the resources that they are familiar using and so the testing, for them, is quite unfair. (Resource teacher, A02)

Exemptions were allowed only in “very rare” circumstances and this was seen as problematic. The testing results for some students might not accurately reflect the students’ growth (*value added*) from the individualized instructional programs. For this reason, the testing was not viewed as entirely fair or appropriate for all students.

Another related concern was teachers’ perceptions (Grade 4/5, M02; Resource teacher, A02) about the stress of the Grade 3 testing and the impact on students’ self-esteem. The Grade 7 teacher, a former Grade 3 and Grade 6 teacher, has administered the EQAO tests several times and described her personal frustration as followings:

The students need to be able to do their best and feel good about coming to school and want to learn forever. That’s what I see in the job. Not, how do I convince them that this week long test is a *good thing* for them, when I know that it is stressing them out. (Grade 7 teacher, A02)

On the other hand, the Resource teacher (A02) described one situation where a Grade 6 student, qualified for exemption, wanted to write the test so that she did not feel singled out from the rest of her classmates. These ethical dilemmas reflected the moral elements inherent in professional practice.

Despite such concerns, Mr. Page and his staff focussed on keeping Northside School was a safe and caring place. The teachers reviewed school performance and made improvements to their programs. The principal reported seeing a positive momentum in school performance reporting a “1 to 2 % climb each year since we have been doing the testing” (Principal, Au02).

The relationship reported between the external and internal instruments and strategies was compatible, perhaps not ideal, but the combined strategies provided oversight to a process that allowed school discretion. In a way, the EQAO testing raised

the visibility of the underlying cause for students' low scores, rallied critical reflection on programming and social issues, and triggered school and community networking.

The Nature of Northside School's Internal Accountability System

The nature of Northside School's internal accountability system is captured here by describing how the school makes an account of its responsibilities to both school and community stakeholders. The EQAO provides oversight to the accounting process by requiring that Northside School report on school (students) performance to the parents and school community. This mandate ensures accountability processes are initiated. How Northside School carries out this duty and fulfills its responsibility to communicate and interpret school performance is now characterized by school response to five recommendations from the EQAO (see Figure H15 for full proofs).

In response to the first recommendation, Mr. Page and the SBDM Committee *consulted with school and community stakeholders* during the annual program review and improvement planning process that follows receipt of the EQAO testing results. Mr. Page discussed school results with the Parent Council and interpreted school performance in terms of all the information on students' achievement. He and the SBDM Committee consulted with district support staff. Committee members also contacted the Signal School network and other community agencies to share ideas about how to address common concerns and issues. This multi-level collaboration was established practice, suggesting a strong response to this first recommendation.

A second EQAO recommendation was the *use of assessment strategies to find out where students are in their learning and what they need to do next*. First, the Northside teachers began to try out different assessment methods to track student

learning. Assessments were done frequently (2 to 3 times a week) and this was reported as supporting formative purposes. Teachers reported pacing their instruction and modifying lesson designs to accommodate learning styles and special needs. Differentiating the lessons to scaffold the intermediate stages of the learning process was reported as common practice (Grade 4 teacher, J02; Grade 6 teacher, J02; Grade 7 teacher, Ma02). Many students at the school had IEPs and this identification required that many instructional programs possess remedial support and modified expectations. In such cases, students were graded based the expectations listed in the IEP. Another reason that diagnostic and formative assessment was common practice is found in the school wide commitment to using the CAT tests. Teachers used these and other test results (e.g., Grade 7 diagnostic math test) to identify students' strengths/weaknesses and to plan regrouping and instructional units.

A third EQAO recommendation, that was common practice at Northside School, was the *identification of school and community factors related to achievement*. As a Signal School, Northside School staff was active in working with parents and community agencies like the city Health Department to identify and support a range of student social and physical needs. The school also looked at transience issues and discussed how to provide continuity in support structures.

The teachers at Northside School communicated about the *purposes of assessment, the achievement expectations and important supports for learning* in two ways. First, teachers reported that they taught learning outcomes through rubrics and discussed evaluative criteria at the beginning of the school year. This put transparency into classroom assessment processes, and teachers reported being surprised at the impact this strategy had on student learning. As the Grade 6 teacher (J02) explained, "I think

rubrics are the reason... I would say that the students are beginning to recognize expectations.” The Grade 4 teacher (J02) found that students exposed to assessment criteria and taught procedures for explaining their reasoning in mathematics “become [became] pretty good at determining what they have been able to communicate clearly and what they had difficulty communicating.”

Second, to keep parents up-to-date on their child’s progress, the Grade 4 teacher began to send *formative and summative assessments* home regularly so that parents knew how the student was managing along the way. She found that:

Along with it are comments written by me, such as *This child will continue to need to work on this* and I will write how it is going to be handled *I will address this during homework club*. If it is something I think parents can handle, then it is mentioned that it is something that could be addressed at home (Grade 4 teacher, J02).

Providing extra support at school was common practice but enlisting parent support with homework was not as consistent at the higher grade levels.

The approach at the Grade 6 level was different. With an emphasis on developing student independence, this teacher encouraged students to meet with him at recess or after school to discuss their results. These small group sessions or one-on-one conferences were tutorials that frequently included some follow-up practice at home. Students recognized that the weekly assessments and marks were recorded and used “at a parent interview when a student is not functioning or in danger of failing a particular topic” but were not sent home for parental review on a regular basis (Grade 6 teacher, J02).

A fifth EQAO recommendation that was established practice at Northside School was the *discussion of achievement and improvement strategies*. The report card was a

formal means but it alone was viewed as inadequate to keep parents informed about student progress. Students with IEPs were evaluated against different expectations and this was often confusing for ESL parents. Teachers appreciated these dilemmas and reported that student progress needed to be explained in terms of real gain and found that parent-teacher interviews were important venues.

The Grade 7 teacher (Ma02) found that her students' parents put a lot of trust in the teachers of Northside School because they did not understand the system. Both the Grade 4(J02) and Grade 6 (J02) teachers explained that meeting with parents was the optimum way to discuss student's achievement. These teachers reported that parent-teacher interviews were important opportunities to explain what the expectations were and what the grades meant. To facilitate parent-teachers interviews, Northside School used a cadre of language liaisons who volunteered as translators so that teachers and parents could discuss student progress in meaningful ways. The Grade 6 teacher (J02) preferred to meet with parents after school on an ad hoc basis, particularly when he was concerned how a student was doing academically. He found that the 20 minutes at a report card interviews was not enough time. In these ways, Northside School worked to develop closer links with the students' families.

Interpreting and communicating expectations and student achievement with teacher, parents and community stakeholders was common practice at Northside School. The unique needs of the student population made this responsibility an important focus for these teachers and administrators. School efforts to bridge the language issues and to share valid interpretations of the testing results were conscious ones. Northside School response to the EQAO recommendations suggests a comprehensive implementation.

This first aim of the EQAO accountability agenda, school accountability to the parents and school community, was fulfilled.

In re-examining the school aims, as defined in the school mission statement, Northside School was acting in ways that reflected the school aim to provide *a safe, challenging and caring environment* and a focus on *human development*. The presence of strategies to address each of the five recommendations indicates the influence of accountability mechanisms to initiative communication between schools and the public. The nature of this response is characterized as explaining and justifying the school instructional program and appears to reflect a high level of accountability.

Summary

The internal accountability system at Northside School was strong and supported school improvement through encouraging responsible decision-making by its staff and by extending the depth and breadth of involvement to teachers, parents and community stakeholders. The EQAO testing agenda and its mechanisms were one environmental influence that impacted the internal accountability. School leadership was also influential because it mobilized supportive actors and allocated resources. Compatibility between the external and internal aims and combined strategies was evident because local needs and priorities could be incorporated into the improvement plans. The EQAO agenda kept literacy and numeracy aims to the forefront, while the district aims supported organizational development. This reveals flexibility in the accountability framework that allowed schools to take into account the social and cultural elements of school communities while at the same time focussing on the business of teaching and learning. The next section of this chapter will examine the influence of the internal accountability system to mediate teachers' assessment practices.

Influence of Northside School's Internal Accountability System on Assessment Practice

The second aim of the Ontario accountability agenda was to promote change in teachers' practices. This section first describes classroom mathematics assessment at Northside School. The second part then captures the collegial dynamics that characterize the influence of the school's internal accountability system on teachers' collective practices.

Classroom Assessment

Classroom assessment at Northside School is reported here in answer to the second research question. Assessment practices were analyzed using four features including: a) the purpose of the assessment, b) the criteria used to select the assessment, c) the method, and d) the feedback provided (see Figure H16). Each feature is now described in sequence.

Purpose

Assessments at Northside School were conducted for diagnostic, formative, summative, student management, and accountability purposes. Each of these five purposes is outlined in order to capture what motivated assessment activities.

Diagnostic assessment was common practice at Northside School and helped the teachers to measure student achievement levels and screen for gaps in knowledge. Teachers reported that pre-assessments helped with unit planning and pre-test results were used to form instructional groupings.

Teachers used observation, quizzes and class discussions to assess students' knowledge and skills. The Grade 4 teacher (J02), in particular, preferred informal oral methods like brainstorming activities but sometimes used "some sort of a written test."

She also mentioned accessing the Grade 3 EQAO tests to support her initial assessment of individual students, particularly those who seemed to be having difficulty with the mathematics lessons. In contrast, the Grade 6 teacher (J02) didn't access his students' Grade 3 EQAO scores but preferred to confer with the previous teacher or double check the Grade 5 mark book that was passed along to him as part of school routine. The Grade 7 teacher (Ma02) reported that she favoured using the district mathematics diagnostic test that all Grade 7 students wrote each September. These results showed the strength and weakness across the different math strand topics and the results were helpful in planning instruction. This teacher commented that she did not use the Grade 6 EQAO results for diagnostic purposes, partly because she questioned the value of scores. Her thinking was the results were limited by the context of the testing (language, length of the test, student anxiety). Northside teachers reported diagnostic assessment as an important part of their practice.

At Northside School *formative assessment* was the main purpose behind the use of journal writing and quizzes as these helped teachers monitor students' progress. Teachers used formative assessments to determine if students understood and applied new concepts or if they needed to be revisited. The Grade 4 teacher (J02) explained that math journaling showed that students "were challenged by the new vocabulary and words were being confused." She reported introducing vocabulary using graphics in response. She explained her habit of thinking, "What it is that I need to revisit? What is it we need to work on as a whole group? or Whom do I need to see at recess time or after school to support along the way?" (Grade 4 teacher, J02).

The Grade 6 teacher (J02) also used math journaling to assess and plan next steps. As he explained, "the nice thing about working in the math journal format or a

quiz where I give them the spaces to work, is that I can usually track down where the students are making the mistakes.” He tried to pinpoint where procedural mistakes were made and when his students made calculations errors he would focus on the procedural steps. He commented that the majority of class time was spent teaching mathematics processes and redirecting students’ attention by using statements like “Let’s get the right answer” and “Let’s do the proper steps to find I,” emphasizing formative feedback in that way.

The Grade 7 math teacher (Ma02) explained that he used the “same assessment instruments as you would use in a junior division, a lot of observation.” This teacher and her team partner used hands-on activities with the students and monitored how students carried out their work. Both teachers were former junior teachers and used this informal approach because it immediately identified the next step.

Many of the students have huge, huge gaps in math knowledge and the only way to find out what the gaps are is to give them things to do, then watch and observe how they act, how far the student can take the activity. (Grade 7 teacher, Ma02)

The Grade 4 teacher explained that assessment continued after a concept was taught. Her students were given many opportunities to practice each concept and she continued “marking on a regular basis what they are doing in their seatwork, although, that is not necessarily used in the final evaluation at the end of each unit” (Grade 4 teacher, J02). How this assessment was tracked is best explained in her own words.

My system is a bit complicated. When the children work on an assignment, even if I just introduced it at the beginning of learning a new concept, I mark it and have a tendency to record it as a percentage in my mark book. That is not ultimately what I will be using because it is not an ultimate assessment strategy. But what I need to know and do is red flag students who I can see are having difficulties...I do that on a regular basis, daily, but then I have a page for their best renditions and [for] the assessment I created for the assessment of progress. And that is recorded on the second page and that

probably happens once or twice a week. But it depends on whether we have [become] stuck on something and need more time because their daily work shows me that they are not quite ready. Then we need to revisit a few things. So it does fluctuate but generally it's once or twice [a week]. (Grade 4 teacher, J02)

Formative assessment as a way to monitor student progress was an integral part of the mathematics instruction in these classrooms. These practices involved teacher-student interactions that focussed on what students needed to learn, what was working, and what needed to be changed to improve student understanding.

Summative assessments are used to decide a student's final grade. Northside teachers frequently used paper and pencil tests for this purpose but also reported using *culminating performance tasks* for some math strands, geometry for example. Teachers reported that key formative assessments were also used to round out the picture of a student's overall achievement. For example, reference to informal demonstrations in other subject activities or conversations on the playground were considered in final grading. Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02) found that together these sources of data provided a fairly comprehensive picture of students' progress.

Northside teachers reported that assessments were also used to motivate students and in this sense used for purposes of *classroom management*. Weekly quizzes and journal assessments were talked about in class and the Grade 4 teacher (J02) commented how "it is highly motivating for them because they get instant feedback and from that they know how they are doing." She found that students began to understand how to prepare for the culminating task and benefited from the supports (homework club, extra practice) put in place to help them. Her perspective was that everyone wants to be successful, so of course, her students participated willingly.

The Grade 6 teacher (J02) found that many of his students had a “low attention span”, while others who struggled with mathematics would “give up easily.” He had students work in groups and to keep them on track, he would use weekly assessment marks and dated comments when dealing with behaviour.

I always tie their behaviour to marks because quite often students forget why they're at school because outside issues are often more important to them. I always try to bring it back to academics and I apply their marks and any notes that I've taken on them and say *Now, let's get back to work.* (Grade 6 teacher, J02)

The Grade 7 teacher also explained that assessments motivated her students, particularly at the beginning of the year when so many students were new to the school.

I will start with the facts and as the students [are] very familiar with [that] and then...I try to build their vocabulary around [this in] the beginning so that they can eventually tackle the more challenging language of grade level work. But I always begin by allowing them to experience success with numbers. To me that's important. (Grade 7 teacher, Ma02)

The Grade 7 team started with basic facts and computations skills in September because success early in the year was important to self-esteem and student status among their peers. These comments reveal that teachers at Northside School used regular assessment to both motivate and manage student behaviour.

Teachers at Northside School also assessed for *accountability purposes*. The provincial EQAO tests conducted in grades 3 and 6 were instrumental in this sense. Students were also assessed for reporting purposes three times each school year, and reports card grades were expected to reflect achievement against the provincial standard. However, the school-based CAT testing that began as a diagnostic instrument also provided a measure of students' performance that was useful when reporting to parents during the school year. If student performance was below grade level in September, approaching the provincial standard during the school year reflected the effectiveness of

the instructional program or if performance did not meet grade level the CAT scores could explain why (Grade 4 teacher, M02).

Assessments were conducted for many purposes at Northside School. The criteria teachers used to select particular instruments and tools for mathematics assessment are described next.

Criteria for Selection

Northside School teachers selected assessment instruments and tools based on several criteria. To illustrate their reasons, the criteria are organized into the five following categories: Origin, relevance in terms of outcome, fairness in terms of scaffolding learning, relevance to the purpose, and economy of use.

One reason given for selecting assessments was the *origin*. Many of the instruments used by Northside teachers were selected from the new mathematics textbook. Some of these assessments were modified, added to with teacher-made questions or computation worksheets. Some teachers reported that they also used parts of the EQAO sample tests and questions from the anchor books. Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02, Grade 7 teacher, Ma02) used the provincial mathematics exemplars that had recently arrived and found these an excellent source of ideas for assessments because they came with rubrics and detailed performance criteria for each level of achievement. Rubrics were widely distributed by the district program support staff and these included rubrics for math journaling.

When the Grade 6 teacher (M02) found himself short of time, he would opt for a streamlined version of the textbook assessments and would “cut back, scale back here and there. Rarely would I substitute unless I felt that we had to move on and I had to get

a concrete result.” He explained that it was difficult to cover units as comprehensively as he would have liked in the time he had.

Another criterion used was related to *economy of use*. Teachers used the assessments embedded in the mathematics program because they saved teacher time and were efficient to use (Grade 6 teacher, J02).

Teachers at Northside School had a third criterion in mind, *methods reflected the intended outcome*, when selecting an assessment. The principal requested that teachers use a variety of assessment methods. Teams met to share best practices and discuss assessment methods suited to different learning outcomes. The methods reported at Northside School included journaling, paper and pencil tests, performance tasks, and text-embedded problems. The journaling focussed on assessing student reasoning and communication. The Grade 4 teacher explained her method for assessing communicating and reasoning.

During group activities, I go around with my clipboard and use a check-off system where I am at each table, listening to what is happening and asking a few questions. I check off on a checklist whether or not they have been able to communicate properly. (Grade 4 teacher, J02)

Teachers reported what they chose depended on the math strand being taught and its expectations. Geometry assessments (Grade 4 teacher, J02) included performance tasks dealing with planning and constructing three-dimensional shapes, for example. These were assessed with checklists and rubrics that looked for problem-solving skill and the ability to identify multiple solutions where appropriate. These multiple methods showed recognition that different types of assessment were more suitable than others to measure specific mathematics outcomes.

Fairness was another criterion when choosing assessments. The Grade 6 teacher (J02) explained his thinking from a classroom management perspective. He found that his students would lose interest and be overwhelmed by a single unit end test. Since they responded better when they had regular feedback, he used several weekly quizzes to assess learning. In doing so, he avoided putting all the weight on a single unit end test. His students had opportunity to see the type of questions he would use for the final test and had ample opportunity to practice at school.

A fifth reason teachers considered was whether *the results suited the intended purpose*. Mathematics journaling was introduced in a workshop and it was a very popular strategy for formative purposes (Grade 7 teacher, Ma02) because the journal entry clearly reflected whether the student understood the math ideas and could explain the question using appropriate terminology. As the Grade 6 teacher explained, he selected journaling “I usually try to get them into the habit of explaining their work, so that they are becoming more familiar with how to detail every step. I work a lot with math journals as an assessment tool” (Grade 6 teacher, J02).

These criteria were sometimes practical, sometimes instructional and sometimes a matter of knowing what suited students’ learning styles. The common elements among these criteria were the importance of using methods that were student friendly and that reflected the type of assessment found in the EQAO tests and provincial exemplars.

Assessment Methods

Teachers at Northside School were required to use a variety of assessment methods. Four common methods described here are classroom recitation, text embedded assessments, paper and pencil tests, and performance tasks.

Classroom recitation included question-answer exchanges, interviews and anecdotal records of conversations. One-on-one conferences and small group discussions were common practice in Northside School classrooms, especially if remediation was needed. The Grade 4 teacher (J02), for example, gave students who had difficulty writing answers the option of doing the journaling as an oral exercise explaining how she asked them to “sit down with me and we go through what they have said, what was their rationale for choosing the operation they chose.”

The Grade 6 teacher (J02) shared how the EQAO emphasis on written communication was problematic when it comes to the ESL students in his class. He often asked students to “demonstrate with manipulatives” or “explain it orally” to determine what they knew. The Grade 7 teacher (Ma02) also reported similar methods and felt that one of her strongest assessment strategies was to observe students and note what they were able to do. Oral questioning was used to scaffold learning and to assess student performance at Northside School and teachers reported that they kept checklists and jot notes to track these measures.

When Northside School replaced the mathematics textbook, they selected one with *text embedded* assessments. These handy resources facilitated frequent assessment as common practice. The math program possessed extensive problem-solving activities and journal prompts as well as quizzes and unit-end tests. Accompanying the text were additional *performance tasks* with rubrics and checklists already prepared. These assessments covered the scope and sequence of the instructional program and contained formative assessments and cumulative activities that reportedly aligned with the provincial expectations. Northside teachers used these materials extensively.

Teachers at Northside School continued to use *traditional paper and pencil tests* supplemented by the use of performance tasks. Although the teachers primarily relied on the assessments materials accompanying the textbook, the content did not match the curriculum in every respect. The Grade 4 teacher (J02) found that she needed to supplement the unit ends tests with her own tests, particularly when it came to computation and basic math skills. The Grade 6 (J02) teacher also reported that when the students required more practice, he would prepare his own paper and pencil tests. In a sample of his assessment instruments, commercially prepared computation sheets were found. This reflected the importance of computational mastery and application of mathematical procedures; his observation was that “open ended explorations did not always work well with his students” (Grade 6 teacher, J02).

This review of assessment methods at Northside School did not include *standardized tests* as these were described elsewhere, for example, the use of the CAT for diagnostic purposes. The variety of assessment methods revealed a commitment by Northside teachers to diversify their assessment. Much of the impetus for this change was directly linked to new textbook and the different methods it employed. Teachers used the new methods but expressed different levels of comfort with the methods. This study did not intend to examine the quality of teacher assessments but will next describe what type of feedback teachers provided to students.

Feedback procedures

Teachers provided feedback to students in several ways. Rubrics were a popular tool because they possessed *qualitative description* and *numerical ratings* that aligned well with the format of the new report cards. The Grade 6 teacher (J02) reported using a *numerical rating scale* (1 to 4) for journal assessments and a *percentage* for tests and

quizzes because his students preferred that format. He explained that “the math journals are the only work that I assess as level 1, 2, 3 or 4 so that the students understand what these concepts mean. I use plus or minuses here and there to give the scores flexibility” (Grade 6 teacher, J02). His feedback was primarily symbolic but his observations resulted in *anecdotal notes* that contributed to how he determined grade and comments for the report cards. *Praise* shared with students was oral in nature and *immediate*.

The Grade 4 teacher (J02) also used *percentages* and *letter grades* in her mark books but preferred to use rubrics for communicating to her students about their progress. She would “write down the areas where there were difficulties in their work, areas of strength where they did well.” She tried to be specific in her *descriptive comments* about what students could do to increase their mark – what they could be working on. Her aim, she explained, was to encourage students or invite them to come for extra help.

Comments made by the teachers at Northside School (Grade 4 teacher, J02; Grade 6 teacher, J02, Grade 7 teacher, A02) revealed that student effort and improvement towards mastery of concepts and processes were valued and recognized. This seems to illustrate a focus on the *process* of learning. The Grade 6 teacher, in working with older students, used formative assessment to show his students what he looked for when marking and this represented feedback about their progress. In his assessment interviews and small group discussions, the feedback focussed on problem-solving strategies and procedures. His aim as he explained it was to move past the calculating accuracy to “point out that they are getting the right steps” (Grade 6 teacher, J02). To support the importance of looking at process, the Grade 6 teacher used group assignments and tracked “how each student help [ed] other people”.

In the classrooms at Northside School, The Grade 4 teacher tried to provide *germane feedback* to her students and she found this to be tricky:

The specific expectations don't necessarily pinpoint precisely what is necessary for each level. For example, if they asked the students to identify geometric shapes, we understand through specific expectations what geometric shapes they are supposed to know. But if the activity involves sorting, then sorting these various shapes into as many categories as you possibly can would create a level 4, in terms of sorting into categories. But what creates a level 3 and what creates a level 2? (Grade 4, J02)

She discussed these points with her team partner and explained that she had had similar conversations with a parent.

This section examined classroom assessment practices at Northside School and characterized the purposes behind student assessment, the instruments and strategies used, and the criteria upon which they were selected. It explored the feedback that teachers provided to student and parents so that an understanding of what teachers were learning to do was pictured.

Dynamics of School Assessment Practice at Northside School

The internal accountability system found at Northside School mediated many of the decisions that individual teachers made at the classroom level. Classroom assessment practices were influenced by the collective decisions of teachers, the activities stipulated in the SIP and policy requirements. The internal accountability system structured these organizational procedures and facilitated professional discussions of assessment. As the second aim of the EQAO agenda was to inform teachers' (assessment) practices, this study now reports on the presence of five characteristics indicative of the nature of assessment communities defined in the literature (see Figure H17). These characteristics

include assessment goals, assessment instruments, assessment methods, personnel involved in assessment, and values.

Assessment Goals at Northside School

The goal of Northside School was to support students attaining the provincial standard of achievement but, in reality, the starting point for many students presented unique challenges. Initially, teachers were concerned about the standard and its carte blanche applicability to the school as a whole, but the professional dialogue during the school review and improvement planning processes turned early resistance to acceptance of standards-based assessment.

Internal accountability system procedures and processes supported assessment goals by focussing teachers' attention on program review and improvement planning. The SIP process led to further diagnostic assessment across the grade levels, dispersing ownership for program evaluation and development to classroom teachers. This led to formal procedures (IPRC, IEP) being followed when students required differentiated or modified instruction based on assessment data. School review guidelines and grading handbooks revealed district support and guidance to bring commonality to teachers' assessment practices. Teachers reported that these interactions facilitated understanding assessment methods and helped them to implement standards-based assessment. These mechanisms guided teachers and mediated the *acceptance and compliance with the external standards*.

Assessment Instruments at Northside School

Northside teachers used the *curricular documents*, *achievement chart* (with level descriptors) and grade level *exemplars* as tools to inform their assessment and grading practices. The Grade 6 teacher accessed the *EQAO anchor booklets* and performance

tests mainly to prepare students for the EQAO tests. Most assessment activities were selected from the approved mathematics textbook and teacher resources, although some supplementary materials were teacher-developed or commercial in origin.

Another instrument that Northside teachers used was the District's handbook for assessment and grading. The handbook outlined assessment principles, recommended multiple assessments detailed how to use rubrics and how to use evaluative criteria to measure student achievement. Northside teachers (Grade 4 teacher, J02; Grade 6 teacher, J02; Grade 7 teacher, A02) reported familiarity with these guidelines and followed the procedures in deciding a report card grades. The principal monitored that teachers followed instructions to report each math strand twice within the school year.

Holistic assessment instruments like *portfolios* that use collective criteria to assess students' achievement were not used at Northside School. Some teachers used portfolios as a method to assess writing development, but this was not a common at the time of the study.

The dissemination of standardized instruments by the province and the district suggests a conscious effort to guide teachers' practices. Use of these common instruments suggests school-wide commitment to consistency in assessment practices and reflected a collective approach that was further informed by the teaming partnerships.

Assessment Methods at Northside School

Northside teachers reported that the administrative staff encouraged them to use of a *variety of assessment strategies*. Diagnostic assessment was established practice and this used standardized test data (CAT, CCAT, EQAO, Grade 7 diagnostic test) as well as teachers' pre-assessments. Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02; Grade

7 teacher, A02) also reported professional development activities that emphasized the importance of *ongoing assessment* and the *formative value* of interviews and anecdotal records was part of team and staff meetings. Best practices such as question-answer interviews, performance tasks and journaling were promoted by teacher leaders. Instruments like rubrics and checklists were also shared and assisted in the application of common analytic and holistic criteria during assessment of achievement. Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02; Grade 7 teacher, A02) explained that demonstrations and journaling supported formative assessment and this helped to scaffold student learning and focus on the *process of learning* as well as the end product. Unit end tests remained a popular culminating method though.

The Grade 4 teacher (J02) explained that she used the *design down process* in planning her mathematics units. She decided upon the culminating activity first, then planned the instructional unit to address the intermediate steps, and then selected the assessments needed along the way. She reported modifying the unit lessons as needed to meet student needs. The Grade 6 teacher (J02) found that when teaching his mathematics program his “strongest support comes from the textbooks, the assessment tools, the black line masters and the practice and testing pages that come with the teacher guide.” As a new teacher, he used the new textbook and was trying out the assessments that came with it. His preferred assessment tools were weekly quizzes and math journaling because he could closely monitor his students’ progress that way.

Northside teachers reported that they assessed students frequently and kept anecdotal records so that students with special needs (ESL, learning disabilities) could be fairly evaluated using a full range of strategies including oral remarks and observed behaviours. Although student and peer assessments were a small part of the data

collection process, the junior and intermediate teachers *owned the responsibility* for evaluating achievement. Final letter grades were reported on the provincial report card for all math strands and written statements qualifying student performance accompanied these letter grades.

Another feature that characterized the assessment practice at Northside School was the *collaboration* between teaching partners in discussing expectations and clarifying assessment criteria. This practice engaged many teachers in collegial dialogue of the specific expectations, the criteria for each level of achievement and how these would be quantified in performance assessments. Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02; Grade 7 teacher, A02) reported that the mentoring, teaming, and divisional team meetings helped to build a *common language of assessment* and provided opportunities to improve their understanding of standards assessment.

Depth and Breadth of Personnel Involvement

At the classroom level, Northside students were taught to use the curriculum expectations (achievement chart) and rubrics to develop their own assessment skills. At the team level, many teachers worked with colleagues to select, develop and critique assessment tasks and evaluative criteria. At the school level, the Resource teacher and SBDM Committee initiated diagnostic testing to inform instructional planning. Also at the school level, the School Review and SIP processes involved conversations between the staff and the parents. Teachers also consulted with district experts and other colleagues while the principal worked with the superintendent. At Northside School, ownership for assessment was inclusive on many levels and exemplified a depth and breadth of involvement that appeared to welcome school and community participation.

Value Systems

Assessment communities are collaborative venues that support reflective practice. This kind of professional learning is enhanced when teachers have opportunities to experiment with new practices and resources at the school and classroom level. The principal and teachers at Northside School valued in-servicing and reported that the staff meetings were used for workshops on best practices and new instructional materials to address the need for professional development. However, teachers did report that time for reflection on practice and critical review of the instructional and assessment materials was limited and that they would welcome more time to engage in professional development activities.

I have gone to a lot of rubric making workshops but I think that one of the things that I would love to be able to have is someone come out and talk to us about exemplars. How they have been used with rubrics because there is still a great bit of subjectivity in using rubrics for assessment and evaluation. I am sort of curious as to whether or not my way of [*sic* thinking is] going to be in keeping with the standards. (Grade 4 teacher, M02)

In terms of what is getting in the way of doing that, basically, the time. The opportunity. I think that is where we, the Ministry, has dropped the ball here with the teachers is in providing the in-service and providing the daylight time for professionals to sit down and develop these skills cooperatively with their peers. There is not enough time provided to them to develop these skills and strategies. (Principal, Au02)

In light of these comments, the introduction of the new curriculum and assessment policies were hindered by what was perceived as a lag in support for teacher development. This view was likely attributable to the restructuring of the educational funding model at the provincial level that resulted in a reduction of Central District's operating budget the previous year. The principal (Au02) reported that there was little budget left for school-wide workshops and that was why staff meetings were now

focusing on professional learning activities. However, internal accountability processes ameliorated this shortcoming by enabling teacher leaders (e.g., Curriculum Contact teacher) to disseminate useful resources and share best practices. Mr. Page and the SBDM Committee kept the pace of change incremental in terms of the SIP goals and activities. As familiarity with the EQAO test and the type of performance tasks it used was achieved, the teachers at Northside School, with the support of District staff, worked to better prepare students for the testing by changing the instructional program and by demystifying the assessment process by teaching about outcomes and sharing criteria with students and parents.

Considering the many changes that had been implemented by his staff the past two years, the principal's perspective was that progress was happening. He particularly noted changes in the teachers' assessment practices.

In the long range plans, we do expect to see the type of assessment tools that will be used. We do ask that a range of assessment strategies be employed. We do have a range of skills within the staff in terms of who is using a more comprehensive approach. Some are more traditional with their assessment strategies. We do talk as a staff about who has an assessment. Where we find we could do more discussion is when we are looking [at] the IEPs and we plan for special needs students and we look for alternative means about doing assessment. It is certainly an area where we need to work on as a staff. (Principal, Au02)

Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02; Grade 7 teacher, A02) were collaborating on assessments, sharing workable assessment strategies, trying to incorporate performance tasks and recognizing the integral role of formative assessment in teaching. Many of these characteristics were likely not new to teacher practice at Northside School but dialogue about standards-based assessment practices was happening. Teachers (Grade 4 teacher, J02; Grade 6 teacher, J02; Grade 7 teacher, A02) reported that they now assessed and measured intermediate steps in the learning process

and found that students were motivated by this type of feedback and seemed more cognizant about the progress that they were making.

Professional collaborations were structured by organizational procedures and school routines but, as the next example demonstrates, revealed a collective ownership for responding to arising concerns.

That [it] happened in one of my math classes this year, where I found that I actually had six students with significant needs. I had enough for an entire remedial group. So what happens [now] with that little group, because it was identified early, we were able to find another teacher with three periods of down time a week, who is [now] doing remediation in math with them. I get them for the three other periods in the cycle. The other teacher and I work together to help that little group move along. (Grade 7 teacher, A02)

This is indicative of a *focus on student needs* by the teachers that extends beyond one's immediate teaching assignment to a school-wide perception of teaching.

Through the School review and SIP processes, the *consultations with parents and community stakeholders* suggest that public participation was valued. These interactions extended ownership for students' learning beyond the responsibility of one teacher. Communication and support for student learning from professional groups and community agencies were enabled and also facilitated a collective response to the needs of the students and their families.

The influence of the internal accountability system to support teachers' assessment practices disclosed how different mechanisms and leadership decisions nurtured collegial interactions characteristic of an assessment community. The mechanisms enabling collegial dialogue and assessment conversations enhanced teachers' practices, the second aim of the EQAO accountability agenda. Collegial assessment practice as opposed to individualistic practice was established at Northside School.

Summary

The internal accountability system at Northside School was strong and proactive; deriving its strength from common aims and compatible strategies between the provincial, district and school level organizations. The mechanisms used to guide teachers' practices at the classroom level were a combination of provincial and district requirements but also included school-generated ones. School leadership facilitated the professional nature of the internal accountability system by extending teachers' ownership for school improvement decisions. Throughout this process, teachers' professional values and commitment focussed on students' needs and learning styles in the renewal of their instructional programs. Professional learning focussed on curriculum renewal through workshops on performance tasks and assessment methods. Commitment to understanding and using provincial standards to assess and support student learning was evidenced in the collegial assessment practices described. Continuous cycles of program review and public consultations reflect a positive response to the dual aims of the EQAO accountability agenda.

CHAPTER 5

INTERPRETATION AND DISCUSSION

This chapter discusses the cross-case analysis for the two schools described in Chapters 3 and 4. The first section interprets the cross-case findings related to the nature of internal accountability systems (Research question 1). This is achieved by examining: a) the elemental components that characterize system configurations (research question 1a), b) the key factors that impact internal systems (research question 1b), c) the relationship between external and internal systems (research question 1c), and d) the effectiveness of these systems to respond to school and community expectations. The first section also compares the influence of internal accountability systems to support the development of teachers' assessment practices (Research question 2). The second section of the chapter summarizes the discussion in terms of the study propositions. These tentative conclusions are presented as suppositions in the sixth chapter, as are the limitations of the study, and the implications this study has for research and for practice.

Cross-Case Findings

This interpretation and discussion is organized by the two main research questions. Following this analysis, the cross-case findings for the second research question are reported in order to determine how the extant internal accountability systems influence teachers' assessment practice.

The intent of this study was to appreciate teachers' perspectives in exploring how accountability aims and processes were interpreted and enacted at the classroom and school level. An administrative perspective is employed in capturing the interpretations

and framing the discussion of accountability systems. The analysis also appreciates the professional dilemmas associated with large-scale assessments (Chrispeels, 1997; Firestone et al., 1998; Grant et al., 1996; Van der Vegt et al., 2001) that are already well-documented in the literature; and therefore explores beyond the initial reactions of the teaching staff in order to capture how accountability policies were interpreted and enacted despite professional concerns and resource-related issues. This cross-case analysis now describes the processes, structures and professional interactions that came into play in both case study schools as teachers and principals accounted for the instructional program and planned for school improvement in spite of the many stresses associated with implementing accountability testing and standards-based assessment policies.

The Nature of Internal Accountability Systems

The research question, *What is the nature of school internal accountability systems in a context of external accountability?*, is discussed by comparing cross-case findings for each of the three subordinate questions. The main question is then answered by encapsulating the overall trends in the findings to determine the configuration and strength of two examples of internal accountability systems located within Ontario's accountability agenda.

Question 1(a) What is the strength of the existing internal accountability systems?

Empirical research has examined internal accountability systems and identified structural characteristics (Newman et al., 1997) and socio-cultural elements (Abelmann & Elmore, 1999; LaPointe, 2000) that comprise such systems. The integrated model presented in the conceptual framework is based on these dimensions. Its five

components are used to contrast the configurations of the internal accountability systems of Southview School and Northside School.

The study findings reveal that both Southview School and Northside School have strong internal accountability systems because each school's internal system, as captured in Chapters 3 and 4, possessed all five components of the integrated model (see Figure 1 in Chapter 1). Newman et al. (1997) reported that strong internal accountability systems are associated with quality instructional programs. Furthermore, since Abelman and Elmore (1999) pointed out that shared expectations (school aims and goals) and guiding mechanisms influenced the nature of the internal accountability, this train of thought suggests that the presence of these components of internal accountability systems requires that teachers' attention be focussed on instructional programs.

It is interesting to note that although both schools possess strong internal systems by definition, some elemental differences within the component parts reveal that distinct configurations of internal accountability are possible. The characteristics reported in the findings are contrasted and provide insights on the degree of horizontal differentiation (Spillane, 1998) in Central District schools. The cross-case findings for each component (see Figure 4) form the basis for the interpretations that follow. This discussion theorizes that a degree of flexibility in the elements of each component at the local school level is possible, and likely preferable (Hargreaves, 1994), in order to foster restructuring and accountability that is responsive to diverse school contexts and conditions. For example, different school communities may gather different types of information on school and student performance that are relevant to local concerns and issues.

In comparing the first component of internal accountability, school *aims*, Southview School and Northside School both identified improvement in student

Components	Southview School	Northside School
Aims	<i>School aims:</i> 1. Improve student achievement through a focus on excellence. 2. Maintain a safe and caring school. 3. Provide an enriched school program. Reports justifying goals and action plans. (professional esteem, teaching reassignment)	<i>School aims:</i> 1. Promote academic and cultural excellence through values education with a global perspective. 2. Develop an action plan to improve students' performance on provincial tests. Reports justifying goals and action plans. (school reputation, professional esteem)
Information on School Performance	<i>Classroom level:</i> Grade 3 CCAT results, previous report cards, classroom assessments, some EQAO results	<i>Classroom level:</i> Grade 3 CCAT results, CAT results, previous report cards, Grade 7 diagnostic math test, classroom assessments, consults with other teachers, some EQAO results
	<i>School:</i> Grade 3 CCAT results, report cards, EQAO results, Grade 7 Diagnostic Math test, standardized testing results done by SERT, EQUIP, socio-economic information	<i>School:</i> Grade 3 CCAT results, report cards, CAT results, EQAO results, Grade 7 Math Diagnostic test, standardized testing results done by SERT, EQUIP, socio-economic information
Standard of Achievement	<i>Classroom level:</i> Achievement chart (the provincial standards), IEP differentiation / modification where appropriate	<i>Classroom level:</i> Achievement chart (the provincial standard), IEP differentiation, modification where appropriate
	<i>School level:</i> Achievement chart (the provincial standards)	<i>School level:</i> Achievement chart (the provincial standard), IEP modification where appropriate
Persons Who Judge Performance	<i>Classroom level:</i> teacher, students	<i>Classroom level:</i> teacher, teaming partners, students
	<i>School level:</i> principal with input from leadership team, teachers	<i>School level:</i> SBDM committee with principal input, teachers
	<i>Community level:</i> parent council, neighbouring school leaders, community parents	<i>Community level:</i> parent council, Signal Schools network, community support groups
	<i>District level:</i> superintendent, support staff input	<i>District level:</i> superintendent, support staff input
Mechanisms	<i>School-based strategies:</i> Strategic plan (principal) SIP- goals, action plans, areas of responsibility School Report, report cards Staff and team meetings (best practices) Networking with local high school Parental sponsorship of programs Textbook replacement (teacher selection) Resources centralized (math manip.) Timetabling change Staff reassignment (team building) Long range plans (new format focussing on program management) Classroom visitations (daily) Annual growth plans linked to TPA Report card cross-referencing Principal's role (as initiator and manager)	<i>School-based strategies:</i> Strategic plan (SBDM committee) SIP- goals, action plans, areas of responsibility School Report, report cards Staff and team meetings (best practices) Networking with Signal schools Networking with Community agencies Textbook replacement (teacher selection) Team teaching, divisional planning Partnerships and mentoring Long range plans (focus on use of assessment toolkit) Annual growth plans Principal's role (as consensus builder)

Figure 4. Cross-case Analysis: Internal Accountability Systems Components.

Note. This display is a composite of the data reported in Chapters 3 and 4 (see Figures H 6, H12).

achievement as important. At each school, this aim was clarified as specific objectives in their respective SIPs, with the former school focussing on writing and mathematics, while the latter on reading and mathematics. Of interest is the fact that both schools flagged enhancing teacher instruction in mathematics as an important goal. In addition, Southview was aiming to develop a caring school environment by providing an enriched educational and artistic program. In contrast, Northside focussed on school climate issues to address the needs of its culturally diverse and transient student population. In fact, the aim to establish a respectful and inclusive school culture was the first priority addressed in its school strategic plan. In light of the socio-cultural needs of the Northside student population, this priority is understandable.

The aims for each school were explicit in their mission statements. These statements represented the vision (the collective voice) of the students, staff and parents about what the school was most accountable for. This collective voice was framed by Hargreaves (1994) as one element associated with a restructuring perspective for educational change that valued professional contributions (see Figure H5).

The second component of internal accountability is the presence of *information on school performance*. Both Southview and Northside Schools collected a variety of data to assist with their program review and school improvement planning. Both schools used socio-economic data, student population characteristics, early learning experiences, grades, and standardized aptitude and achievement tests (especially the EQAO scores). Both schools also gathered qualitative data from school and community stakeholders when feedback on the effectiveness of the school instructional program was shared. The data collection, therefore, possessed quantitative and qualitative information.

The participating schools differed somewhat in their data collection. At Northside School, additional diagnostic information was collected through the CAT testing to screen the highly mobile student population at the beginning of each school year. These tests were used to assess learning needs and plan instructional programs but also helped teachers account for student achievement when speaking to parents about the students' progress in the instructional program throughout the school year (e.g., value-added). At Southview School, the principal cross-referenced individual EQAO results with students' report card grades. She reported that, with this information, she could then decide if the students' performance on the EQAO test was appropriate when contrasted to the students' classroom performance.

The extensive collection of information gathered by both schools reflected the importance of multiple sources of data to conduct a meaningful assessment of school performance (program evaluation). School EQAO results were considered in light of each school's context; its neighbourhood, its student population, and its instructional program. Principals at both schools reported that the EQAO test results and socioeconomic data were important sources of information and received careful analysis.

The third component of an internal accountability system, a *standard of achievement*, was present at both Southview and Northside Schools. The use of the provincial achievement chart to measure students' performance was established practice. Teachers at both schools reported familiarity with the curriculum expectations and acceptance of the provincial standard (level 3 on the achievement chart) for purposes of grading students' achievement. However, several teachers at Northside School perceived that the provincial standard was ambitious for the student population as a whole, so this acceptance was somewhat reluctant. Resource teachers at both schools

reported special needs students required a lot of support to approach this performance target.

Teachers reported concerns about balancing students' need for encouragement and self-esteem with grades that represented objective measures of achievement against the Grade level expectation. This was indicative of a professional orientation that was oriented to supporting student affective needs and appeared to be in conflict with professional responsibility to use formative assessment to recognize the intermediate steps in learning and using this feedback to scaffold new learning. It also discloses a possible tendency to see grades as a means to reward effort and as a form of praise, two aims that focus feedback on the student not on the learning. This is akin to the dilemma of purpose that Monson (1998) describes when teachers' perceive their responsibility to use assessment to support student learning to be in conflict with external accountability demands for standardized measures of student achievement (grading). These concerns suggest that differentiating the formative role of assessment from the evaluation of achievement in grading is not fully comprehended. There may well be confusion about the role of the EQAO testing as an evaluation of student achievement rather than an assessment tool useful for program evaluation.

Teachers at Southview and Northside Schools felt they were knowledgeable about the provincial expectations and achievement standards. They reported that the EQAO tests, performance results and anchors were helpful in this regard. These instruments appeared to be a tangible way to clarify the curriculum expectations and promote teacher awareness of achievement standards while providing feedback on school performance. This does not affirm that all teachers felt the curriculum and achievement standard was appropriate for all students, in fact, many reported that the

expectations were high. The degree of acceptance is discussed in more detail later, but the point to be made here is that the EQAO testing seemed to play an educative role as research on large-scale testing purports (Herman, 1997).

Evidence from the study about the fourth component of internal accountability indicated that input into the *judgment of school performance* came from many sources. Both Southview School and Northside School principals participated in extensive consultation processes with teaching staff, district superintendents, support staff, parent councils, administrative colleagues, and community support groups. In the case of Southview School, the principal also consulted with the principal and parents at the local high school. These assessment conversations (Earl & LeMahieu, 1997) provided feedback on students' performance on the Grade 9 EQAO math test and this informed the school review (evaluation) of the intermediate mathematics program. These conversations enabled community involvement and contributed to the development of collective expectations for the school and its teachers. Similarly, the principal and school leaders at Northside School consulted with the Parent Council, other Signal Schools, and community support groups to address issues of health, transience and second language education. In describing the differences in these assessment conversations, the suburban school targeted academic expectations and instructional improvement while the urban school targeted a broader range of school effectiveness issues including concerns for socio-cultural sensitivity and health, in addition to academic learning.

These findings illustrate the depth and breadth of involvement by teachers, administrators, parents and the community in judging school performance. These assessment conversations were an aim of the EQAO agenda (Earl & LeMahieu, 1997; Earl & Torrance, 2000) and appear representative of a shift from the trust-in-persons to

trust-in-process that Hargreaves (1994) articulates as the challenge for educational change movements. The latter dynamic (trust-in-process) places confidence in communication and consensus building to facilitate meaningful change, while the former appears to rely on individuals acting independently for the good of others. The former he typifies as system reform, while the latter is characterized as systemic restructuring because it includes change between the relationships of persons within the system.

An important aspect of these cross-case findings is that the depth and breadth of communication and consultation among stakeholders was extended in these two elementary schools, and this contrasts with earlier study findings (Earl & Torrance, 2000; Snider, 2001). These studies reported school discussion of EQAO performance was primarily school-based and only involved explaining EQAO results to parents. The consultation described in this study indicates a more dynamic exchange and a more inclusive ownership among the different stakeholders. Mobilizing supportive actors and tapping into available supports and resources (Knapp, 1997) was made possible through these consultations because the explanation of school performance and justification inherent in improvement planning rallied support among diverse stakeholders.

Southview School and Northside School are nested within the same district-provincial context. The cross-case findings reveal that each school was guided and monitored by the same external policy instruments and strategies. These instruments and strategies form the *external mechanisms* that guide and monitor teachers' practice and contribute to the fifth component of an internal accountability system. External mechanisms included the EQAO accountability requirements (SIPs, School Reports), the standardized report card, the TPA procedures, and the Ministry target setting exercise. However, other external mechanisms existed and these were district-level initiatives

such as the School review audit, the school improvement planning (SIP) guidelines and workshop, the guidelines for assessment and grading, the district literacy and numeracy initiatives, the staff development (best practices) workshops, and the school networks.

The School Reports, the SIPs, the report card and the target setting exercise were all activities that required school personnel to make an account of school performance and student performance to the parents and the public. These mechanisms were directly linked to being accountable for the instructional program. This level of accountability required obligatory reporting, explanations of performance, and justification for improvements in the design of the instructional program. According to Leithwood and Earl (2000), this defines a high level of accountability.

Teachers' practice at both schools was also guided and monitored by school organizational procedures and practices. These *internal mechanisms* also contributed to the fifth component of the internal accountability systems. The internal and external mechanisms were not exclusive, but how these related to one another will be discussed later when the relationship between external and internal accountability is further examined. For now, the kind of internal mechanisms operant at Southview and Northside Schools are examined. The internal mechanisms common to both Southview School and Northside School were strategic plans, annual personal growth plans, working staff meetings, new instructional materials, centralization of resources, and long-range plans.

The strategic plans were devised by school principals after the school audit. These plans addressed organizational issues that involved the management of school resources and supervision of the professional staff. At Southview School, the strategic plan was owned by the principal. Her plan addressed concerns with the instructional

program, teaching assignments, and professional dynamics among the staff. She identified the SIP as one of her main responsibilities and took a lead role in the development of the outcome goals and in facilitating the improvement activities. She identified strongly with her role as instructional leader and attended every team meeting to ensure that professional learning opportunities were maximized. She displayed a leadership style that Fullan (1991) identified as being an *initiator*.

In contrast, the strategic plan at Northside School was a collaborative effort mentored by the principal. He integrated the SIP into the SBDM Committee's responsibilities and encouraged this team of administrators and teacher leaders to set long term goals. The SBDM Committee strategic plan was a three-year school improvement initiative. This principal's leadership style is best characterized as *managerial* because he organized resources, supported teachers in their plans, facilitated changes but delegated responsibility among the members of his staff.

Both strategic plans were found to address the same organization and supervision responsibilities but each principal's leadership style impacted on the methods, and therefore, impacted on the nature of the internal accountability systems at both schools. The impact of leadership is further addressed in the discussion of the findings on factors.

Both schools also required that teachers submit an annual growth plan and the principals reported that they asked staff to address one outcome from their respective SIPs. New mathematics programs were being implemented at both schools, this commonality was likely linked to a special one time grant for mathematics textbooks. These were selected from the approved textbook list and, despite not necessarily being popular at first, were better aligned with the curriculum expectations than the textbooks previously used. Working staff meetings were also common to both schools and,

although they were established practice at Northside School, the principal at Southview School had only recently changed the purpose of her staff and team meetings to include a professional learning component and team planning focus.

Long-range plans were also common in both schools, but their structure varied. Long-range plans at Northside needed to detail the type of assessment being used; while at Southview, the principal was intent on reducing the bulky binders down to simple one-page templates for each subject area. The purpose at Northside was to ensure a focus on using assessment to support ongoing feedback to students, while at Southview the purpose was to mentor timely planning and collaboration among teachers. These internal mechanisms focussed on the goals of the SIP on curriculum renewal and were therefore curriculum linked. They enabled the internal accountability system because they monitored teacher practices. The above mechanisms were common to both schools, but three internal mechanisms differed. Two were operant at Southview School and one at Northside School.

At Southview School, two internal mechanisms guided teacher practice directly; teacher reassignment and reorganization of the school timetable. Both strategies were used by the principal and were based on her goal to support the quality of the instructional program. The first action, reassigning teachers to grade levels best suited to their teaching skills, involved an assessment of teaching style and expertise. This resulted in some teachers changing grade levels after being very comfortable with the grade and instructional program they had been teaching for several years. A change in grade level required the development of new instructional units based solely on practical considerations. In a sense, this ensured that teachers spent time developing new curricular units using aligned instructional resources. The second mechanism,

reorganizing the timetable, put mathematics periods in the morning when students were more alert and meant that all intermediate math classes were scheduled at the same time. This change was intended to enable team planning, even team teaching. These actions were taken by the principal in response to feedback from school and community stakeholders about the effectiveness of the intermediate mathematics program. In a sense, these decisions represented a management response to address a concern, but the actions also resulted in issues affecting professional esteem, and as such were a form of extrinsic consequence.

One internal mechanism reported at Northside School only was its annual CAT activity. This strategy was related to school concerns about transience and special needs students. The staff instituted this school-based accountability testing upon the advice of the SBDM Committee. The Committee proposed the idea to the staff and discussed the benefits of completing diagnostic screening in September. This decision brought accountability into the instructional program because it enabled teachers to explain to parents the level of student achievement in September. It enabled teachers to group for instruction to better meet student needs. It also provided a potential measure of instructional effectiveness when it was suggested that a companion test should be administered each spring and student progress assessed. This mechanism brought accountability for the instructional program to each grade level, not just to the Grade 3 and 6 teachers. It indirectly involved obligatory reporting as well as explanation and justification of one's instructional program to teaching colleagues. This school-based decision was intrinsically linked to teacher duty and professional esteem.

In reviewing the findings of this cross-case analysis of the internal systems at Southview and Northside Schools, it is apparent that each school-based system

possessed *common components* but *unique configurations*. The external mechanisms were the same while the internal mechanisms varied in subtle ways. The differences were not tied to the external mechanisms. These differences were a response to the diversity of school needs and demonstrated that flexibility in the mechanisms enabled schools to respond to diverse needs with different strategies. This illustrates that the internal accountability systems were somewhat adaptable to local conditions but sufficiently coherent to address the broader aims of the external accountability agenda. These findings support that a balance between external oversight and internal ownership for accountability is possible (Fullan, 1998).

This sub-section analyzed and qualified the components of two strong internal accountability systems nested within the same provincial and district context. The integrated model facilitated identification of both structural and socio-cultural elements that comprise school accountability systems. The strength of the existing internal accountability systems, the aim of question 1(a), was qualified. The next sub-section describes the cross-case findings related to factors that helped and hindered school-based accountability.

Q 1(b) What factors impact on internal accountability systems?

In analyzing the cross-case findings of factors that impacted on internal accountability systems, two significant policy to practice influences were identified. These included the *external requirements* associated with the *EQAO testing* and the *mobilization of supportive actors by leadership* at the school level. The cross-case analysis (see Figure 5) represents a reduction of the qualitative analysis captured in Chapters 3 and 4 (see Figures H7, H13 for full proofs). The impact of these two factors on the internal systems is outlined next.

Southview School	Northside School
External requirements	
<u>Provincial requirements</u> EQAO - performance-based testing (objective measure) - transparent reporting (awareness, accountability) - program review requirement (assessment, improvement plans) - recommended strategies (communicating & interpreting assessments)	
Mobilization of Supportive Actors	
<u>Leadership</u> • consults with teachers, parents and community leaders (seeks information to inform decisions) • strong personal vision and commitment • first year principal (strategic management plan) • collaborates with administrative team, supports teachers • focussed on student achievement and supportive school (curriculum renewal) • analyzes data to inform decisions (rational decisions) • initiates organizational change (structures teams, schedules) • promotes collegial planning • comprehensive supervision schedule • promotes professional development	<u>Leadership</u> • uses collaboration and consensus-building to make many decisions • school mission statement (collective voice) • experienced principal (SBDM committee plan) • focussed on student and staff needs (school climate, instructional program) • monitors internal accountability (SBDM Committee) but delegates responsibility • seeks information from multiple sources (expertise sought, community agency support) • supports collegial planning, team partnerships, mentors (selects new staff based on school needs) • advocates for teachers and encourages networking • promotes professional learning

Figure 5. Cross-case Analysis: Factors Impacting Internal Accountability Systems.

Note. This display is a composite of the data reported in Chapters 3 and 4 (see Figures H 7, H13).

A review of the cross-case findings reveals that the accountability systems at both Southview School and Northside School were impacted by the *EQAO tests*. These tests informed schools' program reviews and public reporting but also guided teachers' practices. The EQAO test was a significant policy to practice lever that impacted the internal accountability systems because it influenced teachers' actions. The categories used to organize the analysis of this influence include: test and design, test administration, test implementation, and test results.

The EQAO *test design* impacted internal accountability in three ways. Firstly, the test questions were performance-based tasks and covered each of the five mathematics strands. The teachers at both Southview School and Northside School reported that when they first saw them, they knew their students had not been prepared for this type of test. The first impact was curriculum renewal as demonstrated in the use of new mathematics instructional materials and distribution of math anchor books and curriculum exemplars. A second impact was that many teachers used the tests themselves as a source of *prepared curriculum*. The fact that the tests were redesigned with different tasks each year suggests that this was purposive on the part of the EQAO. The study findings suggest that the tests were intended to get teachers' attention and be a source of performance-tasks, and this helped to defray what Corcoran and Goertz (1995) reported as inhibitors to sustainable change - the availability and scope of appropriate instructional materials. The third impact on internal accountability was that the test itself assessed all five math strands and provided clear examples of performance tasks for each one. This informed teacher practice by emphasizing the importance of all strands, not just the fundamentals of computation and problem solving. Unlike the findings in other empirical studies (Chrispeels, 1997) where large-scale tests did not match the

performance-based curriculum expectations, this study supports the proposition that restructuring in the Ontario context was enabled by the introduction of aligned large-scale performance testing. Teachers at both schools reported that the tests did appear to align with the curriculum expectations.

The cross-case findings also indicate that the EQAO *test administration* impacted internal accountability in two important ways. The staff at both Southview and Northside Schools reported that student exemptions for the EQAO tests were rare, despite practical problems providing all the accommodations special needs students normally received and in spite of anxiety exhibited by students who found the tests challenging. For these reasons, teachers reported that the school's results may not accurately reflect the level of achievement for every child. They responded to this dilemma by reorganizing the support services, and, where schools identified significant gaps in students' performance on the five mathematics strands, a renewed focus on early identification of special need students. This triggered applications for special grants that could provide remedial support for weak students. Teachers also reported that they sought accommodations in student programs and in the testing situation itself. The EQAO policy requiring full student participation was likely an attempt to avoid what Herman (2000) refers to as selective inclusion and non-generalizable gains. A second impact related to increased diagnosis of student strengths and weaknesses was that parents were increasingly involved in the development of individualized educational programs. IEPs and program modifications required teachers to consult and collaborate with the resource teacher and parents each term.

The cross-case findings also capture the influence that the EQAO *test implementation* had on internal accountability. The first influence was that teachers felt

responsibility for their students' performance and felt obliged to evaluate students' classroom work against the same achievement standard. The second related influence in internal accountability was that as teachers became more familiar with the testing, they became better informed about the expectations. The teachers at both schools reported that student performance on the tests no longer held surprises, unlike their first experience three years earlier. A third influence of the implementation on internal accountability was that, according to teachers, administrative and community stakeholders focussed on more than the EQAO raw scores in the program review and this perspective was appreciated. School and community members took into account the degree of uncertainty in the scores, sought multiple indicators (other school and non-school factors that determine student achievement) to review the validity of the results, and placed emphasis on improvement (the *value added* aspects of the instructional program), while also avoiding acceptance of low expectations within the school community. This was a function of the way the test results were used in conjunction with other EQAO requirements, specifically the SIPs and recommendations.

It is important to note that *test development* stabilized during the first few years of the EQAO test implementation. The current Grade 6 tests were shorter than the original pilot tests and now included a multiple-choice section. This reduced the time needed to conduct the testing to a one-week period and made the testing experience more manageable according to the teachers.

The cross-case analysis indicates that the EQAO *test results* impacted teacher practice in two important ways. At both Southview and Northside Schools, individual teachers reported that reviewing individual student scores enabled them to gauge their own appraisal of student performance against that of the EQAO. Teachers reported that

the EQAO standards were representative of the curricular expectations. This finding reflected the potential of these tests to be a form of moderation, as suggested by Black (1993) in his extensive review of assessment research. There were limitations to this exercise because the student scores were reported by strand and no comments are written on the actual students' booklets. However, comments from teachers revealed that they did examine student results by strand, considered individual students' performance in some cases, and were usually in agreement with the performance level assigned to their students. This observation was found for teachers at both schools even where different levels of achievement were anticipated.

A second impact of the test results appeared to be the use of the aggregated scores as an external measure of school performance. This meant that the scores informed the program review process when teachers collectively examined the results, looked for trends in the data, and used the analysis to identify instructional priorities. It was at this instance in both schools, where the feedback informed reflection on the instructional program and resulted in an assessment of the quality of the instructional program. The review process illustrates the type of feedback spiral (Kulm & Stuessy, 1991) that ties both program assessment and professional reflection together. It identified strengths and weaknesses in the instructional program and enabled the development of next steps that formed school improvement plans. As the testing became established, the principal at Northside School recognized that some early hypotheses about mathematics achievement led to strategies that were not showing the kind of improvement that his staff was aiming for. He knew it was time to revisit the original hypotheses that informed the development of the three-year SIP. This reflection was program evaluation because it led to a decision to change the school improvement plan.

These cross-case findings appear to indicate that Linn's (2000) suggestions for avoiding unintended negative effects of high stakes testing were functioning within the current EQAO accountability agenda and were being accounted for by the mechanisms of the internal accountability system. Specifically, the *test design* addressed the need for high quality assessment each year, the *test administration* provided safeguards against selective inclusion, and the *test implementation* placed emphasis on growth over time. The *test results* measured school performance against a fixed external standard that enabled informed comparisons but were reviewed in a process that highlighted the importance of multiple indicators. The processes aligning the test data with the *school program review* and *improvement planning* appeared to temper the positive (complacency) and negative (low expectations) effects of the testing. These accountability mechanisms appeared central to the establishment of school-based forums, and revealed how professional staff owned school performance and assessed instructional programs. Such cross-case findings support the propositions in the literature (Cross & Joftus, 1997; Herman, 1997; Linn, 2000) that large-scale tests may a role in accountability frameworks, but this is qualified by the design of the EQAO tests. These tests received teachers' attention, focus that attention on student achievement, and provide models of what is wanted in terms of performance-based tasks to teachers at both elementary schools in this study.

Based on these findings, the impact of EQAO test results as a measure of school performance appeared to be tempered by the compatibility of the external and internal accountability mechanisms to address inappropriate use of the testing data. The testing data provided a single external measure that was co-joined with other sources of

information to inform the program review processes that were a function of the internal accountability system.

A second significant policy to practice factor impacting internal accountability systems was *leadership*. Leadership is a sub-factor of what Knapp (1997) termed *mobilization of supportive actors*. As the EQAO tests informed the internal accountability system, school leadership mobilized the accountability processes by accessing expertise and resources.

Southview and Northside Schools possessed motivated principals who both presented themselves as student-oriented educators. Each principal had earned the respect of his/her staff and was actively engaged in ensuring policy requirements were met at his/her schools. Each oversaw that teachers' practice was guided and monitored using established internal and external mechanisms. Both managed internal accountability systems that possessed the same components but slightly different configurations. As both principals mobilized supportive actors and promoted professional networks, the differences in the configurations appeared to be a function of leadership style. This cross-case analysis now unpacks these leadership styles and the impact of principal actions on school internal accountability.

Fullan (1991) described principal decisions and actions as *action motifs*. His motifs are further categorized as *leadership* functions and *management* for change functions (p. 158). Leadership functions include: articulating a vision, getting shared ownership, and evolutionary planning. Management functions include negotiating demands and resources and coordinating persistent problem-coping. The action motifs for each principal are revisited next, and then the leadership style qualified.

The principal at Southview School was a second year administrator. The previous year, as vice-principal, she conducted the school review audit and *consulted extensively* with teachers, parents, community leaders and district staff. From her review, she developed a *strategic plan* that focussed on student achievement and a supportive, caring school environment. She reported discussing the *school mission statement* with her staff and shared *her vision* of a school where every classroom teacher addresses the whole student in their instructional program.

The Southview School principal instituted several *organizational changes* during her first year as principal. First, she reorganized her staff's *teaching responsibilities* and built new divisional teams based on her assessment of different teachers' strengths and where these strengths would best meet student needs. She intentionally mixed new and experienced teachers on each team. Second, she instituted *working staff and team meetings*. She attended all meetings to ensure a focus on professional development and teacher commitment to improvement plan outcomes. She used staff meetings to introduce *team planning* and formalize *collegial dialogue*. She also *reorganized the timetable* to allow more opportunity for team planning and possible team-teaching among the intermediate math classes. She intended that these interventions build collegiality and share ownership for school change. However, these interventions were too recent to determine if the contrived collegiality at the junior and intermediate levels evolved into genuine collegiality or resulted in balkanization.

In addition to these organizational changes, the Southview School principal initiated a *comprehensive supervision plan* that included daily classroom visits, discussion of long range plans, and accountability for personal growth plans. She provided *funding* for professional development that linked to SIP outcomes.

This principal reported that one of her most important duties was developing and managing the SIP. She worked with her administrative team to examine and analyze all the assessment information on school performance, including the EQAO results. This team shared its analysis with the teachers and then framed the *SIP outcomes* that the school would address in the SIP. The team members suggested some activities but collaborated with teachers in selecting which instructional strategies the school should use to enable the outcomes. The Southview principal showed respect for teachers' time and planned *the pace of change* to what she thought was a manageable level. Teachers were responsible for their instructional program and she wanted them putting their time and energy into *redesigning their mathematics program* and using the new mathematics program materials, including new types of assessments. To facilitate program development she reported introducing a new long-range plan template (one page) that would enable her to liaise with teachers through the year and monitor how they were progressing with their plans. She valued her role as *instructional leader* and was committed to managing change in teachers' practices.

Of particular note is the frequency of the contacts with staff members that was reported by this principal and her teachers. Informal conversations during class visitations and formal discussions during interviews about annual growth plans were interventions to support practical changes and professional reflection. This dialogue provided ongoing feedback and inherent within this professional exchange was consequences that may be intrinsic in terms of professional satisfaction or extrinsic in terms of esteem within the professional community.

The leadership style of Southview School's principal is reflective of Fullan's (1991) *principal as initiator*. This style is likely associated with a professional

management approach to accountability (Leithwood & Earl, 2000) that frames the organization as a whole being accountable, but with more responsibility on the shoulders of the principal. This approach sees the principal accountable to the district officer to whom they report and uses rational administrative procedures to create goal-oriented *effective schools*.

In contrast, the principal at Northside School was an established administrator and had facilitated a number of changes within the school during his four-year tenure. Evidence from the study revealed that a *consensus-building approach* was used to create the *school mission statement*. As follow-up to his school review, the Northside School principal formed a School-Based Decision-Making (SBDM) Committee to review the annual EQAO results and create the SIP. The *school strategic plan* was actually a three-year strategic plan, with a philosophy and stated hypotheses, which was designed to set the direction of the improvement process. The principal encouraged staff to volunteer for the SBDM Committee. He *facilitated* Committee decision-making by providing resources and enabling all members to attend the district SIP workshop. He encouraged Committee *collaboration* with district assessment experts and the superintendent. His leadership was less direct and more *supportive of others*. He did not appear to have a well-defined supervision schedule, but all teachers described him as professionally supportive and *willing to find resources*.

Northside's principal was instrumental in establishing *collegiality* among his staff because he promoted teaching partnerships and arranged for mentors. One influence that was primarily his responsibility was hiring new staff and he reported having specific criteria including expertise in different subjects, assessment and diversity in experiential backgrounds. He facilitated staff participation in school improvement and

organizational change by *delegating responsibility* for professional development activities to teacher leaders and divisional teams. He encouraged his staff to participate in *professional networks*.

This form of leadership focussed on *empowerment* and reflection. All Northside teachers were active participants in the development of the SIP at some level. This meant that the teachers *shared ownership* for the improvement plans. Consequences appeared to be primarily intrinsic in nature because the Northside School principal enabled feedback on teacher practice to be a function of collegial dialogue and discourse. This suggests that reflection was based on teacher's professional values and attitudes. This interpretation is supported by comments about the expected changes in the teacher performance appraisal process that had been recently announced. Staff balkanization seemed to be avoided because teacher concerns were raised, discussed and addressed collectively.

The principal at Northside School manifests a leadership style representative of Fullan's (1991) *principal as manager*. He delegated responsibility and organized resources to facilitate his teachers' instructional programs. He also managed the coordination of SIP activities, checked on implementation, and tried not to impose too much on his teachers at once.

The leadership styles of the principals at Southview School and Northside School seem to have more in common than in contrast. Both principals used leadership strategies that articulated a vision for the school, that enabled shared ownership for the aims and goals of their school, and that established ongoing improvement planning. Both principals used management strategies to put limited resources where they were most needed and coordinated problem-coping so that their teachers could handle the many

demands of system change. These management strategies addressed teacher need for professional learning through leadership that individualized support and that stimulated reflection through modeling professional values and collaborative processes. Both styles embody qualities found within the emerging *professional approach* in Leithwood and colleagues' (1999) accountability scheme (see Figure H5).

According to recent works on transformational leadership (Fullan, 2003; Jantzi & Leithwood, 1996), educational leadership must be rooted in the moral dimension of teaching because these principles and beliefs generate the motivation and intrinsic satisfaction necessary to sustain school development and professional learning. Pivotal to motivating teachers' commitment to change is the principal's ability to create trusting relationships among the school staff (Barnett & McCormick, 2003). The development of relational trust involves more than leadership behaviours such as accessibility, encouragement, resources and recognition, but requires consideration for individual teacher teaching and learning needs. These qualities emanated implicitly within the decisions and actions of both principals. It is prudent to suggest here that since the development of relational trust takes time, the length of principal tenure at Northside School may clarify why the internal accountability system at this school appeared to be better established; while the internal accountability system at Southview School possessed more emergent characteristics. The time frame of this study does not permit further comment on this observation.

School leadership, an important element for mobilizing supportive actors, and EQAO testing, an important policy lever, are disclosed as two significant influences on the school internal accountability systems. The former influenced the internal accountability system by managing the school organization in terms of maximizing

resources and by facilitating teachers' commitment to change. The latter influenced the internal accountability systems by providing data to support program evaluation and by focusing school improvement on student achievement.

These policy to practice factors were environmental influences that helped support school-based accountability. Hindrances to the internal systems were also reported by the teachers and these included the availability of teaching resources and professional development workshops. These issues appeared related to practical matters common in implementation contexts (Chrispeels, 1997) but could likely be explained by the concurrent shift in provincial funding policy that resulted in a reduction of Central District's operating budget. In addition, union sanction of the EQAO-related activities and the College of Teachers recertification policies was another issue at the time this study was conducted. These concerns link to professional qualifications and were addressed in the design of this study. As such, inquiry along this line was not pursued. The next sub-section examines the compatibility between the components of the internal and external accountability systems.

Q1(c) What is the relationship between internal and external accountability systems?

The relationship between EQAO's external system and the two schools' internal systems is captured from a cross-case perspective by first scrutinizing the dynamics between the five components first described in Chapters 3 and 4. This is followed by a recap of school response to reveal trends about the nature of internal accountability systems and how these interactive forums enabled conversations among school and community stakeholders. The cross-case findings are summarized in Figure 6 to facilitate this exegesis of the findings (see Figures H8 & H14 for full proofs).

Provincial Context	District Context	Southview School Context	Northside School Context
Aims & Consequences			
a) assure greater accountability and contribute quality education through assessments and reviews based on objective, reliable and relevant information (EQAO tests scores, EQUIP survey. b) make recommendations for system improvement; monitor DIPs; public reports (scores)	a) direct operational planning to support student learning; inform Board policy and budget decisions; and focus improvement efforts of schools and departments b) explain objectives and decisions to parents and the public; provide explanation to the EQAO (submission of DIP)	a) conduct a school review to identify strengths and weaknesses in school operations; develop a SIP that focuses on district literacy, numeracy and caring schools initiatives appropriate to identified school priorities b) consult, explain and justify program effectiveness, improvement strategies and expectations with district administration, school staff, parents and other community stakeholders	a) implement 3 year strategic plan to improve school climate by development of a mission statement, code of learning and inclusive school culture; implement curriculum renewal 2 year plan to focus on data collection, new resources, staff in-service, parent workshops, SIP reviews that focuses on literacy & numeracy b) consult, explain and justify program effectiveness, improvement strategies and expectations with district & staff, support agencies and parent council
Information on School Performance			
EQAO test results, Data on the learning environment inside & outside school (EQUIP)	Standardized testing (CCAT, EDI, Grade 7 math diagnostic), socio-cultural demographics, program characteristics.	Special needs testing (SERT), classroom assessments and grades, School Review data, demographics, Parent Council and community expectations & priorities.	Special needs testing (SERT), classroom assessments and grades, annual CAT, school review data, Signal Schools Network, Parent Council expectations & priorities, EQUIP & demographic data.
Standard of Achievement			
Provincial achievement chart (standard)	Provincial achievement chart promoted (standards); District guidelines for assessment & grading practices (report cards).	Use of provincial achievement chart (provincial standard), district guidelines, school/community expectations(provincial standard), IEPs where appropriate.	Use of provincial achievement chart (standards), district guidelines, school/community expectations (provincial standard as appropriate), IEPs where appropriate (approx. 50%).
Person(s) Judging School Performance			
EQAO publishes performance results, makes recommendations to stakeholders, provides feedback on DIP.	Superintendent monitors Principal's school audit & strategic plan, including SIP. Quality Assurance personnel provide data for SIPs (statistical analysis, EQUIP, socio-economic & demographic data). District staff publish and distribute the District /School profiles which include DIP & SIP to Board and public.	Principal conducts school review and consults with parents, students and staff. Principal mentors teacher personal growth plans and monitors instructional program daily. Principal consults with Parent Council, superintendent and community leaders on expectations and program. Principal & administrative team consult with district experts to evaluate school performance and set goals for SIP. School staff reviews school performance, owns SIP goals and plan next steps. Principal distributes School Program report and student EQAO results to parents.	Principal forms SBDM Committee and collaborates to develop a 3 year strategic plan that includes SIP Principal requires personal growth plans and long range plans. Principal consults with Parent Council, district superintendent, Signal School Network. SBDM Committee consults with superintendent and district experts, evaluates school performance and sets goals for SIP. School staff reviews school performance, owns SIP goals and plan next steps. Principal distributes School Program report and student EQAO results to parents.

Figure 6. Cross-case Analysis: Nested Context of Accountability Policy.

Note. This display is an interpretation of the composite data reported in Chapters 3 and 4 (see Figures H8, H14).

Mechanisms			
Provincial instrument: - EQAO performance-based testing program Provincial strategies: - district and school EQAO results reported to public - district and school improvement plans required (DIP & SIP) Provincial Feedback: - district and school aggregated scores, individual students scores - recommendations for district and school actions	District instruments: - Board Strategic Plan (2000-2003) - School Review Guide (2000) & audit process - School Improvement Planning Guide (2000) - Statistics and Test Manual (2000) - Assessment and Grading Handbook (1999) - District testing program & school demographic data District strategies: - Supervision of school review process - SIP workshop (expertise & input) - Staff training and development - District literacy, numeracy & safe schools initiatives - Community partnerships District feedback: - Superintendent recommendations for school actions based on school review (implementation of 6 factors rated)	Southview School instruments: - Principal's strategic plan - SIP, School Report, report card - Long range plans (principal's template) Southview School strategies: - Principal conducts school review audit and consults with superintendent and colleagues about strategic plan - Principal revisits school mission statement and consolidates the collective vision - Principal reviews school performance (SIP) and consults with administrative team to set instructional priorities - Principal supervises report card process - team planning promoted (professional dialogue) - professional development (best practices introduced) - working staff meetings (team co-operation and professional learning) - replacement of instructional resources - resource inventories & centralization - primary, junior and intermediate team reorganization (teacher reassignment by principal) - timetabling changes (to enable regrouping & team teaching, increase instructional time) - long range planning (principal monitored, instructional supports focus) - comprehensive supervision (professional growth plans, classroom visitations, team meetings) School feedback: - EQAO results, professional dialogue, principal consultations, and community input - teacher performance appraisal (intrinsic rewards, collegial esteem/sanction)	Northside School instruments: - SBDM Committee strategic plan - SIP, School Report, report card - Long range plans Northside School strategies: - Principal consults with staff, parents and student to develop school vision & mission - School review audit - Principal establishes SBDM Committee as part of strategic plan - SBDM Committee consults with district staff to review school performance, establishes outcomes, and plans SIP with teachers - team-teaching, mentors & divisional meetings - collegial networking promoted (Signal Schools) - best practices promoted - replacement of instructional resources (texts & assessment tools) - resource inventories & centralization - long range planning (principal monitored, assessment strategies a focus) - professional growth plans (upgrading encouraged) School feedback: - EQAO results and feedback from professional reflection, collegial dialogue, district level consultations, and community input - teacher performance appraisals (intrinsic rewards, collegial esteem)

Figure 6 (con.) Cross-case Analysis: Compatibility of Accountability Mechanisms

Note. This display is an interpretation of the composite data reported in Chapters 3 and 4 (see Figures H8, H14).

The first component of an accountability system is the stated aims. The aims of the EQAO accountability program are two fold: first, to ensure that schools are accountable to the public for student achievement and second, to foster enhanced teaching practices. These aims were similar to those stated on Central District's website: to focus school improvement efforts on literacy and numeracy and to explain and justify its district strategic plans and improvement goals. These aims were paralleled in both schools as well. Specifically, both Southview School and Northside School were accountable for reporting and explaining school performance to parents and community stakeholders, and second, both were accountable for conducting program reviews and developing SIPs. These accountability aims were consistent across the three levels of responsibility.

The second component of an accountability system is information on school performance. The EQAO provided a standardized measure of school performance based on the curriculum expectations and provincial achievement chart criteria. It also provided information to schools on the learning environment inside and outside schools through its indicators program (EQUIP). To this information, the Central School District added standardized testing data (CCAT, EDI, Grade 7 math diagnostic scores), school demographic information, and program specific information. Both Southview School and Northside School reported using information from classroom assessment (grades), specialized testing (e.g., measures of aptitude and achievement based on normative measures) and special education documents (IPRC, IEP), but these schools also gathered qualitative data on parent and community expectations. Together these multiple sources provided an extensive collection that took into consideration many factors impacting student achievement and school performance.

The EQAO advocates the use of multiple data sources as the basis for program review, the EQAO performance scores being one. Central District supported this position and both schools reported using a collection of quantitative and qualitative information. This shows compatibility in the policy and practices across the nested context about the appropriate use of the EQAO test results. It is of interest to point out that Northside School collected additional standardized achievement data for all students in every grade each September. This was initiated for diagnostic purposes, but with the commitment to retest each student in the spring; this initiative becomes a form of school-based accountability.

The third accountability component, a standard of achievement, was reported by all participants across the nested context. At the fourth year of EQAO Grade 6 testing, teachers at Southview and Northside Schools reported being familiar with the provincial standard and using it in assessing and measuring student achievement. Some variation in applying the standard at the school level likely existed, based on teacher familiarity with the curriculum and its expectations, their interpretations of the criteria in the EQAO anchor documents and the availability of the provincial exemplars. The exemplar documents were distributed at the time this study was conducted and teachers reported the evaluation criteria were clearly outlined.

For the purposes of EQAO testing, all students in Grades 3 and 6 were evaluated against the provincial standard. This meant students working in differentiated or modified instructional programs were not exempt. So for school accountability purposes, the provincial standard was compatible across the nested context. This was not the case in terms of individual student assessment and grading on report cards when a student possessed an individualized education plan. Educational policy permitted identified

students to be assessed and graded to different expectations if they were screened and found to meet Ministry criteria for special education status. This process of identification involved an IPRC (identification, placement and review committee) meeting where data on student aptitude and achievement identified specific needs. From the determination documents, an IEP was designed to address these needs. This status was frequent among students in Northside School, as evidenced by the size of its 8 member resource, while it was less common at Southview School. Other students also worked with IEPs for a variety of reasons (transience, second language, remediation). These plans detailed how the regular program was differentiated to assist the student address gaps in learning. This type of IEP was quite common at Northside School and meant that student report card grades were written by teachers and interpreted by parents in light of the IEP.

The standard of achievement for measuring school performance was externally set, but school performance was internally interpreted by school stakeholders based on school knowledge of the student population and shared expectations of the school and its community. The expectations reported by the staff at Southview and Northside Schools revealed that stakeholders held somewhat different expectations about how the students would perform on the EQAO tests. This was explained by the different characteristics of the two student populations. It was the responsibility of the principal at each school to build consensus about school expectations (in light of student needs) in consultation with district experts, school professionals, parents and community representatives each year.

The fourth component of an accountability system refers to the persons who judge school performance. This cross-case analysis revealed that policy makers, teachers, parents and community stakeholders were involved in the review of school

performance, but in different capacities and degree of involvement. These trends are now situated in the context of transparent reporting.

At the provincial level, the EQAO publishes provincial, district and school results on its website. The Central School District does the same but puts these results into context by describing its mission, providing details of its community, and giving information about schools' characteristics and programs. The District also reports on its system improvement plans. Both Southview Northside Schools patterned their public reports after the District model. This reporting process led to school and community dialogue via the trustees and various community special interest groups. Public involvement was systemic and brought qualitative dimensions to the school reviews.

Central District contributed to the review of school performance. This was done through the SIP workshop in a comprehensive manner. District support staff provided statistical data and superintendents gave suggestions on how to use the results to review the trends, contrast performance from year to year, and interpret recent school performance in light of all the information available. These sessions were an opportunity for collaboration between district and school personnel, and both Southview School and Northside School attended these workshops. Parent Council representatives had the option to attend but reportedly had not to date.

The district also provided input into the review process through the district literacy and numeracy initiatives. These initiatives required each school to set one outcome in each of these areas in their SIP. Southview School and Northside School both reported selecting literacy and numeracy outcomes that were based on the review findings. District-level consultations with the public and community leaders resulted in a third district wide initiative in support of school climate. The former initiatives are

understood in terms of the EQAO accountability aims, but the third initiative was in response to public concerns raised at School Board meetings.

As a result of its public consultation process, the Central District established a special category of schools designated as Signal Schools. These schools possessed high-risk populations and received special support including additional funding from the district and access to support programs from community agencies. Such actions could be interpreted as a judgement of school performance and an indication that despite differences in social and economic factors, district and community leaders expected schools to aim for high expectations because the additional supports were provided.

At the school level, principals and teachers participated in the review of school performance. At Southview School, the administrative team attended the district workshop, while at Northside School all members of its SBDM Committee were involved. The structure of the two school committees differed with Northside's team including teachers from all teaching teams.

The cross-case findings disclosed that both principals meet with the Parent Councils to discuss and interpret the EQAO results in light of the multiple sources of information. This represents public participation in the process. The Southview principal also consulted with staff and parents from the local high school. She sought understanding of community expectations and learned about community concerns. The Northside principal sent representatives to the Signal School Network meetings and used these opportunities to discuss issues related to student needs, school expectations and program improvement. These consultations were a source of qualitative feedback and fostered the development of shared expectations between both schools and their communities.

A comparison of teacher involvement in the review process indicated subtle differences between Southview School and Northside School. At Southview School, the responsibility for reviewing school performance was mainly handled by the principal with support from her administrative team. The principal consulted with her staff at a staff meeting and later individually met with the Grade 3 and 6 teachers. She engaged in assessment conversations to develop shared expectations, but also to inform her evaluation of the instructional program. From this review, she identified the outcomes for the SIP and the specifics of the school improvement plan activities were fleshed out with teacher involvement.

The school performance review process was somewhat different at Northside School. The SBDM Committee collaborated with district experts, the superintendent, and the principal. When the EQAO results were analysed, the analysis was shared with the staff as a whole. Some further discussion was done at divisional team meetings and a sense of collective expectations developed. The principal contributed input from his discussions with the Parent Council. The SBDM Committee was collectively responsible for the assessment of school performance and decided on the SIP outcomes to be addressed. The Committee then developed the specifics of the SIP based on the input from the teachers during team meetings. The principal facilitated these processes.

In summarizing the cross-case findings related to decision-making responsibility (depth) and organizational involvement (breadth) across the nested systems, compatibility was noted in two ways. First, each system level consulted with the public stakeholders most directly aligned with their responsibilities and duties. These consultations involved discussions of EQAO results and interpretations based on multiple sources of information. Second, transparent reporting to the public led to an

expanded role for school and community stakeholders. Evidence from both case studies supports this observation. The depth and breadth of persons involved in the assessment of Southview and Northside School performance was inclusive because the public was involved in clarifying the expectations. As the involvement focussed on improvement plans, the ownership among the teachers, parents and community members was extended. These findings lend credence to Hargreaves' (1994) position that decision-making in restructuring systems is not centralized nor decentralized but dispersed among many levels. The cross-case findings provide evidence of this type of responsible decision-making in action.

These findings also suggest that the relationship between the external and internal accountability systems enabled conversations between school and community stakeholders. Consultations with district staff were also extensive and this reveals that these schools were engaged in multiple levels of communication. The nature of these conversations focussed on school performance, what it meant in terms of assessing the instructional program, and what it meant in terms of school improvement plans. These case study findings lend support to the Earl and LeMahieu (1997) position that *assessment conversations* can balance the need for coherent standards with the need for school sense-making (shared expectations) in judging school performance.

The fifth component between the external and internal systems relates to the mechanisms used guide and monitor teachers' practices. The cross-case findings disclosed common instruments and strategies were reported at both Southview School and Northside School. The compatibility of the relationship between the internal and external mechanisms is described next.

External mechanisms common to both schools included the EQAO tests, the public reports, and the requirement schools develop improvement plans. Additional district-level instruments and strategies were also found, revealing the role district policymakers played in school-level accountability. District strategies included: school audits, district testing, literacy and numeracy initiatives, staff development programs, and professional supervision. Tangible district instruments included four manuals (SIP guidelines, School Audit guidelines, assessment and grading guidelines, statistical manual), a diagnostic mathematics test, and software technology to facilitate a common format for the public reports. Other district strategies included professional networks (e.g., Signal School Network) and school-community partnerships (e.g., language liaisons).

At the provincial level, the design of the EQAO performance tests complemented the SIP process because the assessment modeled concrete examples of integrated performance tasks, demanded complex thinking, were curriculum-based, used a variety of response formats and were a potential instructional resource as new assessments were developed each year. Test results were an external measure of school performance that reportedly contributed to a review of the instruction program. The public reports of school performance triggered consultations and collaborations at both Southview School and Northside School. These mechanisms appeared to connect provincial oversight to school-based ownership for school improvement.

At the district level, the staff supported school operations, professional development, and program management as part of its system improvement plan. As this plan encompassed the provincial requirement for a DIP, a compatibility with the EQAO accountability strategies was apparent. The cross-case findings reveal how the District

guided the management of school operations through the school audit process, the strategic management plans, and the oversight of the SIP process. The first of these processes introduced an element of school-community input, the second an element of managerial control, and the third an element of professional responsibility. The purpose of these district mechanisms seemed compatible with those of the EQAO, while also supporting school-based learning and professional ownership for the program evaluation. Since oversight and guidance were the intent, this also represented district support for responsible decision-making.

At the school level, Southview and Northside Schools possessed similar accountability mechanisms. These included external requirements such as the principal's strategic plan, the SIP, the School Report, the standardized report card, and long range plans. The principals and teachers also reported other strategies such as: working staff and team meetings, team-teaching and mentoring partnerships, replacement of instructional materials, centralization of resources, report card requirements, school-based professional learning and annual growth plans.

Differences between the internal mechanisms at Southview School and Northside School related to how responsibility for school decisions and activities was shared and the nature of professional interactions. Differences were noted in staff involvement in the program review, teacher supervision, professional interactions, staff reassignment and school reorganization. These mechanisms related to principal management functions were operant in both schools, but the nature of the principal's interventions revealed a degree of variation (Connelly, 1993; Fullan, 1991).

Leadership differences were minor except in two areas. The first difference related to how the principal at Southview School managed teaching reassignments to

address concerns about the intermediate mathematics program. She made a rational management decision and reassigned the teacher was despite resistance on by the teacher involved.

The second difference in leadership was revealed in how Northside School established its internal accountability testing. In response to staff interest to identify special needs students and arrange remedial support, the SBDM Committee proposed testing all students with the CAT each September. All staff committed to this testing. The SBDM Committee, at the suggestion of the resources teacher, also recommended that all teachers re-test their class in the spring to have a measure of student growth over the school year. The principal supported the idea and facilitated the process by providing the necessary resources. Northside staff supported the idea because it would enable a measure of program effectiveness (value-added) each year, however, teachers may not fully comprehend the degree of individual accountability arising from annual post-testing of student growth. This collective decision alludes to school-wide expectations and reflects the confidence that the teachers felt about the effectiveness of their instructional program.

While both schools' principals supported internal accountability, the cross-case findings note subtle differences in how teachers received feedback in these two instances. These differences appear to represent distinct *approaches to professional accountability*. The nature of feedback at Southview resulted in a loss of professional esteem, while at the same time represented a commitment to putting the right person in the right spot (Fullan, 2003). This was in a sense providing individual support to the teacher being reassigned – a development orientation associated with transformational leadership (Geijsel et al., 2003) – as the reassigned teacher later reported being happy

with the change. The Northside instance emanates an emphasis on vision-building and intellectual stimulation, two aspects inherent in transformational leadership, that supported collegial interaction and ownership for school development through professional learning. At Northside, feedback on teaching effectiveness came as the CAT testing fostered group goals and stimulated teacher inquiry into clarifying individual student's needs. This diagnostic testing led to a refinement of the instructional programs through student regrouping and differentiated instruction, while simultaneously encouraging high expectations.

The consequences inherent in the accountability systems at both schools were a function of the principal's leadership style. This style was described at length in Chapters 3 and 4, but the manner in which decision-making authority was shared (delegated) and the nature of the principal interventions appears to reflect District use of effective schools philosophy and strategies. Both styles seem associated with an accountability approach possessing a professional practice orientation (Leithwood & Earl, 2000) that bears further unpacking.

Both instances can be captured as demonstrating aspects of transformational leadership. In facilitating teacher commitment to proactive change, principals must make decisions based on the needs of the school and its teachers. These decisions may represent a preferred leadership style and perception of embodied accountability, but another interpretation is possible. Consider the school audit process as a means for the principal to identify school strengths and needs on several planes. One plane would be teacher capacity and fit within the needs of the instructional program. The development of the principal's strategic plan would consider possible strategies to address concerns and issues. The strategies selected may well reflect the aspect of leadership deemed most

suitable to the situation. This leads one to ponder if principal motifs (Fullan, 1991) represent personal style or if they actually expose a decision-making process about the appropriate balance between the leadership function and the management for change function necessary to foster school improvement. If principals must understand the contextual constraints placed on schools, as Barnett et al. (2003) suggest, being able to adjust one's leadership behaviours to address the school needs is key to enabling proactive school improvement.

The mechanisms in the cross-case findings and the leadership decisions associated with them lend further support for the proposition that accountability in these contexts is process-oriented as opposed to product oriented. In a sense this may be interpreted as evidence of change as process as opposed to change as compliance.

In summarizing the relationship between the external and internal mechanisms in these two schools, several points are worthy of comment. First, the instruments and strategies at each level of the system appear to be compatible, supporting the proposition that combined strategies (Hargreaves, 1994) support strong internal accountability and foster change in teachers' practices. Second, the reporting of EQAO results fostered school and community consultations. This lends support for what Earl and LeMahieu (1997) advocated in their position that assessment conversations can foster school and community dialogue that strengthens school accountability and community support. Third, parent and community input provided qualitative feedback about school performance and built consensus on school aims. It contributed to the program review and extended ownership for school performance by engaging stakeholders in the process of problem-solving. Such community support may well broaden the conception of school-community accountability. Fourth, the design of the accountability instruments

(e.g., EQAO results, SIP) initiated a continuous cycle of program review, while also permitting a degree of local discretion on what actions should be undertaken to best meet students' needs. These findings suggest that concerns (Linn, 2000) associated with large-scale tests have been addressed. Fifth, district support played an important role in enabling accountability processes because the school audit triggered school-specific strategic planning and the district guidelines for improvement planning and assessment built consensus about expectations and processes; such actions reflect aspects of both effective management and system restructuring (Hargreaves, 1994).

The cross-case findings of the relationship between internal and external mechanisms supports the premise that combined strategies were established and that these processes appeared to balance external guidance (oversight) with school responsibility for accountability and improvement.

This interpretation of the cross-case findings now focuses on analyzing the effectiveness of the internal accountability systems to respond to first of EQAO's dual aims- that schools be accountable to school and community stakeholders (Earl & Torrance, 2000). What follows is a summary of *how* schools made an account of their program to parents and community in response to the first guiding question,

Q1 What is the nature of school internal accountability systems in a context of external accountability?

One aim of the EQAO accountability agenda was to assure greater accountability through large-scale assessments and transparent reporting. This aim was articulated in specific recommendations for communicating and interpreting strategies that schools should employ (EQAO, 1997). The nature of an internal accountability system is therefore reflected in how the school accounts for the effectiveness of its instructional

program. The cross-case findings now describe how Southview School and Northside School responded to five EQAO recommendations related to communicating about school performance and student achievement to school and community stakeholders. The schools' responses reported in Figure 7 represent a reduction of the information described in Chapters 3 and 4 (see Figures H9 and H15 for full proofs). The five recommendations include: a) consulting with teachers, parents and the community to establish and review actions plans; b) using classroom strategies to find out where students are in their learning and what they need to do next; c) identifying school and community factors related to student achievement; d) communicating about the purposes of assessment, the expectations, and quality of student work; and lastly, e) discussing achievement and improvement strategies. School responses are now outlined for each of the five recommendations.

The cross-case findings for the first recommendation, *consulting with teachers, parents and the community to establish and review actions plans*, indicate that consultations with parents and community leaders was established practice at both Southview School and Northside School. The School audit and program review processes were instrumental in initiating these types of interactions. The SIP process also structured this type of dialogue. The Southview School principal discussed the EQAO results and improvement plans with the Parent Council and the local high school stakeholders. The SBDM Committee at Northside School collaborated with several groups during the review of school performance and the development of its SIP. These school-community conversations were a means to build shared expectations and rally all available resources to support school plans.

General Recommendations <i>Specific activities for Schools</i>	Southview School Response	Northside School Response
Communicating & interpreting 2. Schools consult with teachers, parents and the community to establish actions plans and review them on a regular basis. 5. Classroom assessment strategies to find out where students are in their learning and what they need to do next. 35. Identify school and community factors related to student achievement. 36. Communicate about the purposes of assessment, the expectations for students, the quality of student work and the ways that the school and the home provide support for learning. 37. Discuss achievement and improvement strategies.	2. implementation is established, the process and plan are effective, improvements are continuing. [SIP process, School Review audit, School Report, consultation with parent Council and Community stakeholders] 5. practice is common and variety of strategies being implemented, difficulties have been identified and improvement initiatives are being carried out. [performance based tasks, formative assessment in textbook series, target setting, preparing for the test] 35. school audit completed and improvement strategies in place for instructional program and school operations, implementation is being monitored by principal and superintendent, results are documented and communicated to stakeholders. [School Review audit, SIP, EQUIP and socio-economic information gathered from consultations, Community input] 36. principal and teachers using assessment toolkits, difficulties identified and improvement initiatives being carried out to involve parents in extra practice/assessment discussions, student taught expectations and evaluative criteria. [student assessment (curriculum expectations and criteria taught), homework help, rubrics and checklists, Parent Council dialogue] 37. high levels of implementation and effectiveness are evident, implementation is monitored, results documented. [SIP, report card, parent-teacher interviews, IEP, consultations with community stakeholders, School Report]	2. implementation complete, the process and plan are effective, improvements are continuing. [SIP, School Review process, School Report] 5. a process and plan for implementing regular diagnostic testing for all students in place, difficulties identified and improvement initiatives are being carried out. [CAT testing, extra support from resource teacher(s), performance based tasks, formative assessments new instructional materials, target setting, preparing for the test] 35. school review completed and EQAO and other information screened, improvement strategies in place for inclusive school climate, renewal of instructional program, Signal School identification. [School Review, SIP, School Report, Signal School Network, community partnerships] 36. principal and teachers using assessment toolkits, difficulties identified and improvement initiatives being carried out. Student taught expectations and criteria. Communication with parents seen as important. [rubrics, checklists, formative assessment, parent-teacher interviews, language liaison program, Parent Council dialogue] 37. implementation of SBDM Committee plan monitored by principal and superintendent, results are documented and communicated to stakeholders, high levels of implementation and effectiveness are evident, implementation is monitored, results documented. [IEP, report cards, parent-teacher interviews, collegial assessment practice, professional learning, consultations with Parent Council and supportive health agencies]

Figure 7. Cross-case Analysis: Schools Response to Accountability

Note. This display is an interpretation of the composite data reported in Chapters 3 and 4 (see Figures H9, H15).

The cross-case findings for the second recommendation, *using classroom strategies to find out where students are in their learning and what they need to do next*, was common practice at Southview School and established practice at Northside School. Both schools introduced new mathematics instructional materials that aligned better with the curriculum expectations and possessed embedded assessment tools. All teachers reported using these materials. All teachers used informal observations of students' progress to gauge student learning. Planned formative assessment was more common among the early junior grades than among the older elementary grades.

Northside School had a large special needs student population and this meant that many students had IEPs. These plans were updated each term and student progress was monitored carefully. As a result, the teachers at Northside School were in the habit of using formal assessments two or three times a week. Northside School also instituted the CAT each September to measure students' achievement in language and mathematics. These diagnostic tests and strategies helped the teachers monitor student learning and plan instruction.

The third recommendation, *identifying school and community factors related to student achievement*, was established practice at both schools. It was an inherent part of the district's School audit and annual review of school performance. The Central School District gathered district testing data and demographic information to supplement the EQAO data. The consultations that each school principal had with Parent Councils and other community stakeholders also provided qualitative information that helped to frame the factors affecting student achievement.

The fourth recommendation, *communicate about the purposes of assessment, the expectations, and quality of student work and the ways that the school and the home*

provide support for learning, was common practice at both Southview School and Northside School. Teachers at both schools reported teaching students about curriculum expectations, discussing the assessment criteria in rubrics, and explaining the performance expected by the provincial standard. Home support for learning was limited at Northside School because of second language issues.

The fifth recommendation, *discuss achievement and improvement strategies*, was established practice at both Southview School and Northside School. The standardized report cards, the IEPs, and parent-teachers interviews were the means by which teachers communicated with parents about student achievement and plans for improvement. At the school level, the program review process initiated consultations between school staff and parents that informed the development of school improvement plans. The Southview School principal also met with local high school teachers and parents to dialogue on arising issues. At Northside School, the consultation was much more extensive. It involved district support staff, the Signal School network, and various community agencies. This was directly related to the high-risk characteristics of its largely immigrant population. The School Report was another tangible way this information was disseminated.

These five recommended strategies represent how Southview School and Northside School fulfilled their responsibilities to report and communicate on school and student performance. In both cases, these schools accounted for their program and their practice by obligatory reporting, explaining and justifying the instructional program through assessment conversations with students, parents, and school and community stakeholders. Through these communications strategies, both schools made an account for their instructional programs.

These cross-case findings indicate that the operant internal accountability systems enabled both Southview School and Northside School to respond to the first aim of the EQAO agenda, public demand for accountability. The influence of these internal accountability systems is examined next in order to describe their influence on enhancing teachers' practices, a second aim of the EQAO accountability agenda.

The Influence of Internal Accountability on Teachers' Practices

In the literature, internal accountability systems are linked to organizational capacity. They constitute a forum for school improvement and change because they join the organizational structures that guide and monitor teacher practices (mechanisms such as school procedures, requirements) with the collegial interactions that comprise opportunities for professional learning. The case study findings reveal that both elementary schools possessed strong internal accountability systems, and that these internal systems possessed similar configurations. The findings suggest these systems facilitated school-community communication about school programs. Furthermore, the findings support the idea that assessment conversations helped to frame aims and goals. The cross-case analysis now examines the influence these school-based forums had on teachers' assessment practices in order to appreciate the characteristics of collegial assessment practice and to explore how internal accountability systems facilitated standards-based assessment (assessment community). How these systems influenced teachers' practices is analyzed to learn whether the second aim of EQAO's agenda was realized.

The findings in Chapters 3 and 4 characterized classroom assessment and described teachers' assessment practices in terms of individualistic versus collegial

qualities (see Figure 3, Chapter 1). The second purpose of this study was to learn how changes in teachers' assessment practice are influenced in restructuring systems where strong internal accountability is valued. This intent was framed within the following research question.

Q2 What is the mediating influence of internal accountability, if any, on teachers' assessment practice?

Cross-case findings revealed that the extant internal accountability systems did influence teachers' assessment practice in both case studies. Descriptions of classroom assessment and teachers' assessment practices are detailed in Chapter 3 and Chapter 4. The cross-case findings presented in Figure 8 present a reduction of the data (see Figures H 10, H11, H16, H 17 for full proofs) and qualify the degree to which collegial assessment characteristics were present in the two elementary schools at the time of the study. For the purposes of cross-case analysis, the degree to which each school exhibited the characteristics of collegial practice associated with assessment communities was allocated to a category (emerging, extant, and established) to clarify the level of the implementation reported. If a characteristic feature was not observed, reported or documented in the case study findings, it is reported as not evident. *Emergent* refers to the characteristics reported in some teacher practice. *Extant* refers to a characteristic reported to be in common practice or consistently used by several (but not all) teachers at the school. Lastly, *established* refers to characteristics rooted in all teachers' (school) practices. The five characteristics of an assessment community used to organize the following interpretation and discussion include: a) acceptance and compliance with the goals of standards-based assessment, b) use of diverse assessment instruments based on

Characteristics of School Assessment	Southview	Northside
Goals: Acceptance and compliance (versus reluctance)	established practice	extant practice ^a
Instruments: <u>Achievement level</u> Determined through collective discussion of criteria versus independent interpretation	established practice ^c	established practice
<u>Portfolio</u> Holistic criteria in use for some, all subjects	emerging practice ^b	emerging practice
<u>Use of exemplars</u> Uses provincial exemplars and anchor (criteria) versus commercial products	established practice	established practice
Assessment Methods: <u>Collection</u> - planned versus spontaneous - embedded in instruction - toolkit approach (multiple strategies) - a focus on process goals - diagnostic, formative & summative	- extant practice - extant practice - extant practice - extant practice - extant practice	- extant practice - established practice - established practice - extant practice - established practice
<u>Assessment literate</u> know and use assessment terms to communicate	extant	extant
<u>Moderation</u> collaborative development of tasks (performance) and tools (rubrics)	emerging practice	extant practice
Personnel: <u>School-wide involvement in standards approach</u> - regular review of school performance - professional dialogue on assessment	- established practice - established practice	- established practice - established practice
<u>Assessment as collective enterprise</u> - students involved in assessment - parents and colleagues consulted	- extant practice - established practice	- extant practice - established practice
Values: <u>Teachers value assessment processes</u> - assessment as instruction - articulates benefits of student assessment - articulates benefits of learning process	- extant practice - emerging practice - extant practice	- established practice - extant practice - established practice
<u>Teachers seeks support and resources from district and community</u> - workshops and resources - parents, parent council, community organizations and professional groups	- established practice - established practice	- established practice - established practice
<u>Teacher classroom assessment focuses on student</u> - formative assessment focuses on process learning - assessments address learning styles - results used for planning	- extant practice - extant practice - extant practice	- established practice - established practice - established practice

Figure 8. Cross-case Analysis: School (Teachers') Assessment Practices.

Note. This display is an interpretation of the composite data reported in Chapters 3 and 4 (see Figures 11, H17).

^a emerging practice (reported by some teachers), ^b extant practice (reported to be used by several but not all teachers),

^c established practice (reported to used by all teachers).

specified criteria, c) use of multiple methods of assessment, d) organizational involvement, and e) professional value system (Hall & Harding, 2002).

The first characteristic of an assessment community is the *acceptance and compliance of the goals* with standards assessment policy. This refers to the schools' acceptance of the provincial standard for assessing student performance. The goals of standards-based assessment were *established* practice at both schools. Some individual teachers initially expressed reluctance about the ambitious nature of the provincial standard because they were concerned about it being appropriate for all students. However, even these teachers agreed that having detailed curriculum expectations and an identified standard of achievement was good for knowing what they should be planning and teaching. Teachers reported that administrators, parents and local communities did recognize the differences among student populations and that multiple sources of information were used in assessing school performance during the annual program review.

This acceptance of standards policy was supported by the internal accountability systems in several ways. These school-based forums used EQAO's testing results to gauge school performance in light of the multiple factors that impact students' achievement within the context of each school's unique environment. Through school improvement planning, principals, teachers and parents examined provincial expectations and analyzed school performance in terms of the specified criteria. The use of the provincial achievement for program review was also accompanied by school-wide activities that examined anchor documents and provincial exemplars to better appreciate what the achievement levels meant in terms of specific concepts, procedures, applications and expressive language. These actions were carried over to the assessment

and grading practices used in the instructional program. Curriculum-linked textbooks with embedded assessment also promoted school acceptance and compliance with standards policy. All teachers reported that reviewing the EQAO results and developing the SIP were instrumental in informing professional discussions and supporting professional learning. These assessment conversations helped to build comfort and confidence applying the new expectations and communicating them to students and parents.

Initial teacher resistance to using the provincial achievement standard to evaluate students' progress was not unlike the dilemmas that Monson (1998) identified in her empirical study of performance-based assessment. Monson found that teachers were concerned about student self-esteem when student work was assessed against a standard (dilemma of role). Teachers in this study also wanted to encourage students and reward effort, so the fact that standards assessment policy required judgement of performance in addition to guidance of performance was problematic. A second dilemma that teachers reported was concern that all students were to be assessed against the standard as demonstrated in the EQAO testing and not the school expectations (dilemma of purpose). Such dilemmas may reveal a period of adjustment where responsibility for student assessment and grading to a standardized standard was emphasized over possible earlier perspectives that assessment was feedback on effort and commitment.

Since the EQAO data was used as a measure of school performance, not individual student achievement, the dilemma of purpose was addressed because students working on modified or differentiated programs were graded based on the expectations in their formal IEPs. Professional values did motivate all teachers to prepare their students to write the EQAO performance tests, and this meant that all non-identified

students were exposed to a more comprehensive instructional program as teachers reported curriculum renewal and exposure to a variety of performance tasks and assessment methods. The dilemma of purpose was addressed through school accountability mechanisms that featured professional learning (e.g., SIP, best practices, team meetings).

In both case study schools, different perspectives about accountability testing were found; however, collegial dialogue among teachers and principals brought issues into the open where they could be addressed. For example, Northside teachers felt more students should have IEPs with individual needs and accommodations clarified. Such individual plans were a means to address gaps in student learning as well as identify students with significant learning needs. This issue resulted in the resource teacher conducting a workshop for the staff about the IEP process, when and when not to IEP, and what an IEP should involve. Early and ongoing monitoring of student achievement became a common practice and teachers at every grade level consistently conducted diagnostic as well as formative assessments. The cross-case findings indicated that teachers, who initially possessed some reservations about standards policy on student morale, were noting improvement in student exposure to the curriculum and improved student skills. Such comments revealed teachers recognized some benefits arising from the accountability testing.

The second characteristic of an assessment community is the use of *instruments* designed to clarify criteria and standards. Assessment communities select instruments that employ specific criteria to define a consistent standard of achievement. These instruments moderate or temper teachers' understanding and application of the standard when assessing and grading students.

Teachers at both schools were knowledgeable of provincial curriculum and expectations. They reported using the provincial achievement chart, anchor books, and exemplars to develop their understanding of the curricular expectations and inform their assessment decisions. Teachers also reported that these provincial documents provided assessment criteria that differentiated the four levels of achievement and these were discussed with their colleagues. Junior level teachers used the assessment instruments and tools that came with the approved mathematics textbooks because these aligned with the type of performance tasks modeled on the EQAO tests and used embedded assessment with common criteria. They were encouraged to do so likely because this insured a degree of consistency in evaluating student performance.

In addition, the provincial report card and its software program were reported as helpful in assisting with reporting student achievement. The district assessment and grading guidelines determined the frequency that each of the five mathematics strands were reported, and advised on how to plan assessments and assign final grades. These external instruments contributed to consistency in teachers' assessment practices by promoting the differentiation of the mathematics topics into five strands and by sketching out strategies to evaluate the criteria for each component within a strand. The use of common criteria was established at both schools because instruments that modeled them were provided and teachers reported using them.

Teachers did not use portfolios as an assessment instrument in mathematics in either school. They did collect information from several assessment activities (i.e., weekly quizzes) during a unit to inform their evaluation of student achievement. As such, this element of collegial assessment practice was only emerging in school practice in the form of writing folders. It is suggested that some teachers within each school were

scouting out the strategy as a literacy initiative, and this instrument may well receive more attention in future. The Southview School principal, for example, mentioned that life portfolios were a future goal in her strategic plan and this seems to imply an appreciation for an assessment instrument that uses holistic criteria.

Cross-case findings support the view that teachers' use of collective criteria to assess student performance was established practice at both schools. This was likely a function of instruments provided by the province and the district for teachers to use. However, it is important to clarify that teacher dialogue and sense-making of these instruments through school-based accountability strategies such as textbook replacement, team planning and best practices presentation brought teachers together to discuss their interpretations and understanding of achievement criteria. These discussions and discourse reinforced teachers' use of collective criteria.

The third characteristic of collegial assessment representative of an assessment community is the use of multiple *assessment methods*. This characteristic includes planning a collection of assessment activities, communicating about assessment using a common language, and moderating understanding of what different performances should exhibit. These practices associated with collegial assessment were extant at both Southview and Northside Schools.

Cross-case analysis indicates that assessments were frequently planned in advance, often because they were embedded in the textbook materials. This practice was common in both schools, but not consistent to all teachers as some teachers reported making adjustments to their math assessments during and after instruction. Advance planning was supported by school workshops advocating design-down planning. Teachers described design-down planning as beginning with a review of the curriculum,

selecting key expectations, and pulling together resources (including assessments) to teach the unit. Variations between individual teacher practice and grade level strategies were reported.

The findings further reveal that teachers were assessing students more frequently. Several reasons were given for this. Some teachers used the assessments for formative purposes, while others reported that frequent quizzes motivated students. Other teachers articulated how feedback contributed to student learning by emphasizing the process of learning. Many teachers stated that it was school policy because principals asked teachers to use a variety of assessments strategies so that different learning styles were taken into account. At Northside School, this practice was monitored by means of long-range plans. Southview School used staff meetings to promote these assessment practices, but assessment tools were not monitored per say. This focus on frequency introduced an element of fairness into the assessment process and facilitated teachers looking at assessment in non-traditional ways.

Teachers at both elementary schools reported that assessments were used for diagnostic, formative and summative purposes, although this was not found to be consistent for all teachers at Southview School. Most formative assessment at the grade 6 and 7 levels, for example, was oral in nature and provided feedback that tended to be praise or recognition of effort. However, all teachers did report teaching students about achievement expectations using rubrics and checklists; and this reveals the intent to communicate assessment criteria and demystify the assessment process. Sharing criteria that explains progress and points to next steps reflected growing appreciation of the formative dimensions of assessment. Teachers at both schools recognized the subjective aspect of assessment and sought to use objective criteria to evaluate achievement, but

they readily acknowledged that the decision on the final grade was tempered by student effort when holistic performance was borderline between two levels.

This study does not address the quality of the teachers' assessments, but the internal accountability systems at both schools encouraged teachers to experiment with new assessment methods. This, in and of itself, is a first step in professional learning. This experimentation was facilitated by SIP goals, collegial planning, and school-based professional development. All teachers participating in the study indicated an interest in developing their understanding of assessment, and this suggests that in using the new assessments (and the new criteria) they were becoming sensitized to the complexity of standards-based assessment and the multiple purposes for which students may be assessed.

The cross-case findings illuminated differences in the degree of teacher collaboration on the development of assessment tasks at the two schools. The moderation of assessment, the collective discussion of criteria and appropriate exemplars, is associated with assessment communities and was found to be common practice at Northside School where teachers worked with grade-level partners, mentors and supportive divisional teams. These collaborative structures provided opportunities for discussions of expectations, design of suitable assessment tools, and selection of appropriate criteria. Similar collaboration was reported among primary teachers at Southview School, but not at the junior and intermediate levels. The Southview School principal was hoping to facilitate this type of cooperation between her teachers through team planning at staff meetings. The importance of collegial dialogue and debate is well documented in the literature (Darling-Hammond, 1990; Hargreaves, 1996; Little, 1993), and the decision of the Southview principal to formally structure these dynamics among

her junior and intermediate teachers was purposive. This management decision created a context that, according to Peterson and colleagues (1996), supports professional learning and change in teachers' ideas, beliefs, and practices. Such strategies indicate how the school-based accountability facilitated moderated practice.

The *fourth characteristic* that contributes to the development of assessment communities related to the *personnel* participating in the assessment enterprise. In dynamic assessment communities, school-wide involvement in the review of school performance is one aspect and collective involvement of colleagues, parents and students in assessment a second aspect.

The cross-case findings show that the breadth of teacher involvement in program evaluation was expanded at both case study schools. This was enabled by the ongoing SIP process that involved analyses of school performance information, identification of program strengths and weaknesses, and participation in curriculum renewal. All teachers were involved in improvement planning, although responsibilities differed somewhat between the two schools. School and community stakeholders also participated in consultations that clarified community expectations and priorities. School personnel also collaborated with district experts and supportive community agencies. A depth and breadth of personnel in school performance reviews was established practice.

The cross-case findings also uncover the degree to which assessment was a collective enterprise involving students at the classroom level and parents and colleagues at the classroom and school levels. Teachers reported that they discussed the provincial achievement chart and taught assessment criteria to their students. Teachers used rubrics and checklists to communicate with students and their parents about expectations and student progress. Students received feedback from teachers in a range of formats.

Asking students to do self/peer assessments were not reported as common practice, but instances were reported in mathematics at the grade 7 level at Southview School (comment on group presentations) and at the grade 4 level in both case study schools (comment on how they felt about their learning). Student self-assessment was primarily indirect through reflection on teacher feedback, performance descriptions on rubrics and checklists, and grades shared with students.

Assessment conversations were reported on many levels within the case study schools and involved administrators, teachers, students and parents. These conversations were inherent in the accountability processes across both cases because a) school reports and student report cards provided information; b) SIPs were developed cooperatively; c) collaborative structures (e.g., team meetings) invited professional dialogue; d) instructional resources facilitated curriculum renewal, and e) supervision (long range plans, personal growth plans) provided individualized support for teacher learning. These mechanisms stimulated assessment conversations on many levels, influenced teacher practice, and promoted a collective effort by principals, teachers and other school stakeholders.

School leadership was instrumental in facilitating school and community participation in assessment conversations. Leadership is driven by values and values are the fifth characteristic that Hall and Harding (2002) use to identify teachers practices in an assessment community.

The *fifth characteristic* of an assessment community is reflected in the *professional values* that underlay teachers' practices. The cross-case analysis indicates commonalities appear to exist between the values reported by the case study participants

in both schools. The common elements to be discussed include valuing assessment processes, seeking support and resources, and focussing on student learning processes.

The cross-case findings indicated that both Southview and Northside Schools were engaging in activities that explored new assessment methods and resulted in several assessments of different formats being used to assess student achievement. Both schools were using mathematics textbooks with embedded assessments and both school discussed best practices and shared assessment tools at team meetings.

The degree of commitment to assessment processes was more pronounced at Northside School. The cross-case analysis disclosed this emphasis on the development and use of multiple methods and response formats among Northside teachers' practice. The prevalence of second language issues at Northside meant that teachers were expected to outline assessment methods in their long range plans – revealing how closely this aspect of the instruction program was monitored. They were encouraged to gather anecdotal evidence of student learning in innovative ways (e.g., recess conversations). Team teaching was common practice and teachers shared responsibility for unit planning and assessment design. Teachers also divulged the custom of passing on marking records to receiving teachers as a way to build continuity and awareness of students' progress. The collective decision to conduct CAT each September highlights the degree to which the staff valued assessment processes.

In examining assessment processes, the cross-case findings suggest an *added-on* orientation to teacher assessment (Spillane, 1998). This may well be a natural feature during a transitional stage- until familiarity and confidence with performance tasks and multiple response formats are fully established, it would seem that teachers continue to use the type of assessments that they are familiar with. The unit end test was not

avoided, but reliance on a single assessment of learning outcomes was no longer common practice at either school.

Assessment data informed teacher decisions on instructional groupings and program differentiation needs. Teachers worked with resource teachers to support remedial instruction. Teachers reported that students welcomed extra help and were motivated when next steps were described. The findings also show that teachers sought parent support for student learning. Informal communication between home and school were emerging as important interfaces to communicate about student progress so that early identification of gaps in student learning initiated extra practice at school (homework clubs) and at home (drill and review). Teachers reported that it was common practice to ask parents to assist with homework and to review previously taught concepts.

The cross-case analysis also highlighted the type of support sought from district personnel and community groups. It appeared to begin with the program review and improvement planning. Schools attended district workshops and identified needed resources and in-service through the activities listed in their SIPs. The findings capture the networking among principals, their superintendents and schools within their local area. Community agencies were asked to play a role in handling health and social issues in support of teachers and needy families. Principals facilitated Parent Council support in school activities and in information sharing at both Southview and Northside Schools.

The cross-case analysis shows some support for the position that classroom assessment was beginning to focus on raising student awareness. Teachers reported that students benefited from regular assessments because the comments encouraged students and the feedback identified next steps for learning. Findings revealed that attention was

being given to measuring the intermediate stages of student learning in many classrooms and the use of formative assessment to scaffold student learning was a conscious action reported by many teachers. It would be optimistic to claim that this function of assessment permeated all teachers' practices at both schools, but the potential of formative assessment to bring attention to the process of learning was present to some degree.

The principal at Northside School made the use of multiple assessment strategies a component of teacher long-range plans. This student-oriented focus was established practice because it was grounded in, firstly, the need to address student learning styles and language skills, and secondly, the importance of focussing on the learning process. At the time of this study, assessment for student feedback was common (extant) practice at Southview School, but the diversity of assessment strategies and use of results for planning was not as extensive. Trends in the analysis illustrate that assessment was done for purposes other than grading by teachers at both schools. Teachers reported that they monitored student progress with frequent assessments and this was a more formalized process than in the past. Teacher-student assessment conversations (feedback) were common practice but likely included praise and supportive messages as much as descriptions of needed improvements. Rubrics and comments on assignments and tests were sent home for parental review. These practices suggest that classroom assessment practice was less an individualistic exercise but more interactive because these tools helped students and parents appreciate assignment expectations and talk about the criteria defining good performances.

The cross-case analysis of the five characteristics of assessment community suggests that both schools possessed several qualities associated with *assessment*

communities, although not to the same degree. Although not every aspect of the five characteristics was established practice, the overall composition of teachers' assessment practices indicates the presence of collegial assessment practice. The cross-case analysis points to the proposition that internal accountability systems supported teacher learning about standards policy. These school-based forums appeared to rely on collaborative strategies and school-based accountability mechanisms to facilitate professional dialogue and to ensure opportunities for assessment conversations.

The cross-case findings lend support for the position that internal accountability systems are proactive in encouraging collegial assessment practices within a policy context based on standards philosophy. Accountability mechanisms seemed to promote assessment conversations among teachers, principals and school stakeholders. These conversations, facilitated by school leaders, enabled reflective practice among teachers, fostered collective expectations among participants, and likely expanded a sense of shared ownership for student learning. The cross-case findings capture the characteristics of an assessment community and elucidate the school-based processes that mediated teachers' assessment practices, and in doing so, lends support for the position that the second aim of the EQAO accountability agenda, enhanced teacher practice, was being realized.

Discussion

The purpose of this study was to examine the nature of school internal accountability systems and to explore their influence on teachers' assessment practices. This chapter analyzed the cross-case findings for two elementary schools situated within a restructuring provincial system. This chapter first reviewed the components of the

schools' internal accountability systems, the key factors that impacted them, the relationship between the internal and external systems, and how each school accounted for its instructional program. The chapter next analyzed teachers' assessment practices in order to note the degree to which each aspect of collegial assessment was present. The influence of school-based mechanisms to support the implementation of standards-based policies and assessment practices emerged. To facilitate conclusions for this collective case study, the findings of the cross-case analysis are further discussed below.

This exploratory study of internal accountability sought to characterize the nature of the existing internal accountability systems within two high capacity elementary schools. Based on the cross-case findings, these school-level systems supported the dual aims of EQAO's accountability agenda by facilitating inclusive assessment processes and developmental strategies that guided program reviews and monitored improvement planning. These systems were forums in which school personnel engaged in dialogue with colleagues, support staff, parents and community leaders about school aims, school performance and school expectations. These assessment conversations led to collegial (school-level) and collaborative (multi-level) activity in support of the instructional program and student learning.

The first point of discussion is that the findings of this collective case study support the position that strong internal accountability systems are present in the two elementary schools studied. The findings illustrate that all five components of the integrated model (see Figure 1, Chapter 1) outlined in the conceptual framework are present in both schools. The structural and socio-cultural dimensions of the model were found to be instrumental in framing the policy-to-practice connections and informing on

how each school organization was accountable for its performance. Furthermore, how the internal accountability system enables change in teachers' practices was illuminated.

Although teachers identified the EQAO tests as assessments of curriculum expectations and models of performance tasks, they did not appear to conceptualize accountability as a school-based forum where responsibility for program evaluation could be professional by design but inclusive by process. Teacher participants described the elements comprising each component (aims, the achievement standard, information on performance, persons who make judgements, mechanisms), but did not articulate a relationship between them. Nor did teachers seem to recognize the degree of responsibility that schools possessed in terms of how the accounting unfolded. The cross-case findings capture how teachers were fully engaged in learning how to implement curriculum renewal and deal with increased public attention to school performance and activities. They were aware of the public reporting and the improvement planning but seemed unsure what impacts might come from cursory comparisons of school scores by the public. Teachers at both schools reported that school and community stakeholders were supportive and had reasonable expectations for school performance. This situation was anticipated and influenced the decision to use a backward mapping approach in the study design to capture a sense of what was influencing teachers' practices. Teachers were taking ownership for improving their teaching (assessment) practices but not all seemed to fully comfortable with advocating their position with the public at large. This responsibility was the principals' and the degree to which each principal used school leaders in the process was a function of the school context. Teachers worked primarily with students and parents, principals with groups of parents and community stakeholders.

Principals were more knowledgeable about accountability requirements and processes. Both principals were very clear about their responsibilities for managing such matters. This suggests that the overarching provincial accountability framework, while being implemented, was not yet formalized to the point where issues of responsibility and authority were clearly established and understood by many school and community stakeholders. Again, this was anticipated in the design of this exploratory study and the reason interviews were conducted with representatives from the district and provincial organizations. Policymakers at each level understood what they were required to do and articulated their responsibilities. Delineation between responsibilities was reported, for example, the EQAO representative explained that performance scores were provided to stakeholders and districts and schools were required to provide public reports and prepare improvement plans. The EQAO made recommendations to districts, but did not stipulate how each district or school should carry out its responsibilities.

The engagement of public stakeholders in review processes was reported and principals explained that feedback was useful in informing the improvement plans. Assessment conversations with parents and members of the community about next steps and needed supports were happening, but a closer look at these interactive discussions warrants further study.

The integrated conceptual model, presented as a tentative construct, was useful in capturing the participants' perspectives as well as describing the accountability instruments and strategies. The constructs of this model were further developed as the initial data collection and analysis led to further study of the literature. As the study progressed and analysis of the ongoing data collection unfolded, additional reading unpacked the complexity of policy to practice connections in the case studies.

The integrated model was instrumental in identifying the school structures, accountability instruments and professional interactions that operated to facilitate school change. These processes may best be organized by some thoughts shared by Fullan (1991), who spoke to the challenge of reform as being more than the implementation of single innovations. He flagged the need to move beyond innovations that were merely *first-order changes* designed to improve the effectiveness of current practices, and advocated *second-order changes* that seek to alter the fundamental ways in which organizations are structured. Connelly (1993) categorized change as renewal, reform or restructuring. Renewal appears to align with first order change, while reform alters existing procedures rules and roles to enable organizations to adapt the way they function to satisfy different requirements. Reform therefore seems to incorporate both first and second order change. To fine tune what differentiates reform from restructuring, restructuring appears to extend beyond organizational adaptation because it involves activity designed to change fundamental assumptions, practices and relationships both within the organization and between the organization and its external environment. Such diverse intentions must be both intimidating and expensive to schools and districts. How systemic changes are interpreted and enacted is a function of the whole organization with each nested level revisiting their roles and responsibilities. As such, representatives from the multiple layers of responsibility were interviewed with the intent to capture each group's conception of what schools are responsible for and to whom schools are accountable. These descriptions were analyzed with the integrated model and the cross-case findings suggest that although school-based accountability was the motivating aim the EQAO agenda, the design of EQAO's program was also

facilitating ownership and informed decision-making by educators at each level of the system.

Changing educational systems in fundamental ways is a daunting task, but this exploratory study does provide a sense of the policies, instruments and strategies that shaped the nested accountability context in Ontario and describes how two Ontario schools handled their responsibility for reporting to the public and for improving teachers' practices.

The schools participating in this study were engaged in both first and second order changes. School-level mechanisms that triggered improvement in instructional practices (first order change) included the new textbooks, the EQAO tests and anchor books, and the provincial exemplar documents. These approved resources modeled and articulated the expectations of the new curriculum. School staff was also engaged in second order change as they began to work in cooperative teams to support professional learning and develop collegial assessment practices. Northside School appeared to have these school-level processes well established, while the principal at Southview School was actively implementing structured opportunities for professional dialogue and collaboration.

Second order change was further extended in both schools when school leaders facilitated collaborative networking with colleagues, district support personnel, parents and community agencies. Through the school audits, program reviews, and improvement planning processes, a fundamental shift about responsibility for school performance and students' achievement was expanded to include administrators, community support agencies, professional organizations and parent groups by the very fact that engaging in the review of school programs triggered collective ownership for the many factors that

influence student achievement. These conversations engaged parents and community stakeholders in consultation about improvement planning; it helped to rally services and supports for the school and its students. Many of these dynamics were well established in one school and emergent in the other, however the case study findings do capture reported changes in the practices and relationships at the schools. The integrated model was instrumental in illuminating decisions and activities associated with change processes and helped to characterize the nature of the accountability environment in Ontario.

This discussion now looks at another point emerging from the cross-case findings related to ideas and assumptions that seemed to frame the accountability agenda. The EQAO agenda (Earl & Torrance, 2000) was premised on ideas that arose from the standards movement and from a perspective that valued teacher professionalism. The current accountability agenda was influential in the development and support of strong internal accountability systems in the two elementary schools participating in this study. This finding reinforces Newman and colleagues (1997) position that strong internal accountability is associated with high capacity schools because capacity was a controlled variable in the study design and both schools had strong systems. The findings reveal that these schools were successfully engaged in ongoing reviews of school performance and were establishing school-community assessment conversations on many levels. These observations add clarification to earlier works (Earl & Torrance, 2000; Snider, 2001) that reported little assessment conversation beyond teacher-teacher and parent-teacher discussion of EQAO results by capturing a more dynamic and inclusive interaction between schools and community members in Central School District. Findings in this study reveal how of school-community

discussions about school performance and program development were carried out. The depth and breadth of persons involved in accounting processes was found to be more comprehensive and inclusive in these two schools than previously reported.

Another point emerging from this study of internal accountability was that the internal accountability systems possessed common components and, within each component, common key elements. The commonality of the key elements implies compatibility between provincial, district and school aims and strategies. Some variation in school-level mechanisms was reported and these were explained in terms of differences in local school and community priorities.

Principals used the school audit and program review processes to build shared expectations and set school priorities. They used common instruments like SIPs, new instructional resources, planning templates, professional growth plans, and assessment guidelines to monitor teachers' practice. They used common strategies such as collegial dialogue, supervision, and instructional leadership to guide professional learning. Subtle differences in school configurations appeared to be a function of how principal's decision-making in response to the goals of his/her strategic plan. These differences related to what were perceived priorities and how responsibilities were delegated and to whom. The established principal tended to delegate responsibilities to teacher leaders and to support mentoring relationships, while the beginning principal focussed on supporting curriculum renewal and structuring opportunities for collegial cooperation. These approaches are likely explained in term of where teacher time was most valued at the time – developing collegial assessment or developing instructional programs. These leadership decisions seemed reflective of school priorities as stated in SIPs and mission statements but were also implicit within the strategic management plan.

The third discussion point relates to the professional approach to accountability that was described in the case study schools. The strategic plans, although developed by each principal, were based on extensive consultations with both school and community stakeholders. The principals' strategic plans involved what Fullan (1991) called action motifs. These included both leadership and management decisions, and were informed by knowledge of the students, the teaching staff, and the collective expectations of the school and community stakeholders.

The cross-case findings suggest that the subtle differences in the configurations of the internal accountability systems were a function of the principal's leadership style and reflected his/ her perspective about educator's professional role and responsibilities. The Southview School principal, as an *initiator*, used a *management (effective schools) approach* that focussed accountability on initiating both first order and second order change. The Northside School principal, as a *manager*, used a *professional practice (standards) approach* that focussed accountability on restructuring relationships between people. This appears to emphasize an extension of second order change, but does not imply that first order change was not happening. One interpretation acknowledges Connelly's (1993) position that change focuses on altering "existing procedures, rules, and requirements to enable the organization to adapt the way it functions to new circumstances or requirements" (p. 8). A second interpretation creates a picture that lends support to Hargeaves' (1994) conception that restructuring focuses on changing the roles and relationships between people (i.e. collaboration, collegiality, facilitation, definition of community). Both school-based approaches described above fall into a professional classification within Leithwood and colleague's (1999) accountability schema because both internal accountability systems reflect an innate acceptance of

professional values rooted in moral responsibility. These findings indicate that EQAO's accountability model fostered a professional orientation at the school level.

Whether differences in management decisions were a function of personal style or a management response to a developmental stage related to policy implementation was not determinable within the timeframe and scope of this study. These findings do lead to interesting queries about the reasons for the variation in approach, and foster speculation about principal selection and district reassignment processes. Questions like: What training is required by principal candidates? What leadership criteria are considered in the school reassignment? What role does the superintendent play in facilitating strategic management plans? These questions ponder the intermediate role of districts in managing the accountability agenda and facilitating school-based systems.

A fourth point of discussion now turns attention to the impact of the EQAO tests as a standards-based policy lever. The cross-case findings add insights to earlier reports (Earl & Torrance, 2000; Firestone et al., 1998; Linn, 1993) by specifying the influence that EQAO's performance tests had on the school internal accountability systems. The testing program supplied external measures of school performance that triggered program reviews and helped schools to identify instructional priorities by way of improvement planning. It facilitated goal-setting, for example, to improve program effectiveness by selecting textbooks that addressed all five strands mathematics strands. The review of all school data, the EQAO test results being only one, informed school improvement planning. The tests also provided a motivation for ongoing review because yearly scores were felt to be a tentative indication of the effectiveness of improvement plans, keeping in mind the differences inherent in each cohort of students.

Another impact of these tests was the role they seemed to play in triggering teacher motivation to identify students with learning problems early in the school year. This was especially significant at Northside School, where transience and language issues were concerns. School-based accountability testing and an increased application of IEPs characterized the school's response that led to more individualized educational plans and an emphasis on rallying special education support. At Southview School, where EQAO results were considered very good, the impact on the internal accountability system was to focus greater urgency on curriculum renewal through the purchase of a textbook featuring performance tasks and an assessment toolkit (multiple methods). The study findings reveal that these impacts were significant and proactive.

In addition, the cross-case findings suggest that the negative effectives associated with large-scale performance-based testing (Linn, 2000) were addressed in the design and implementation of the EQAO accountability program. Transparent reporting meant that school and district results are shared with the public, but the nature of the communication and interpretation of the results was to be facilitated by school personnel. Discussions of school performance and reviews of the instruction program were explained and justified by school staff based on a collection of information and principals explained that these assessment conversations allowed persons to consider all factors that influenced student achievement. The principals reported that parents wanted to know what was being done in term of school improvement and offered support where they could. At Southview, support was tangible in terms of volunteers in the classroom and funding for enrichment activities. At Northside, this support was translated as confidence in the school plans to bring healthcare agencies and cultural liaisons into the school to support students and their families. This focus on improvement came from the

process-orientation of the EQAO program that made the testing data only one of multiple sources of information on school performance. The school programs were reviewed with an eye to improvement and to incorporate changes that would address students learning more effectively- a value added orientation. No consequences related to sanctions or rewards to individual teachers or school staff were reported.

The fifth point of discussion speaks to the relationship between the internal accountability systems and the external provincial accountability system. The aims, instruments and strategies were found to be compatible because the external system mandated improvement-oriented processes that required school and community stakeholders to review information and engage in explanations and justifications of instructional programs. These activities prompted responsible decision-making through a menu approach to policy implementation. In other words, provincial requirements focussed on public reporting and improvement planning that was based at the school level. The EQAO reviewed plans and made suggestions but the content and processes were local decisions. Accountability tests provided substantive input in terms of objective measures of school performance but the EQAO recommended other measures of school performance be used as well. District staff through strategic plans (DIP) provided guidance and channelled resources to support SIP goals activities. These external supports refocused school-level attention on students learning (achievement), but seem respectful that school-based decision-making (SIP, collegial assessment, etc.) was critical to designing effective instructional programs that addressed local student needs and community priorities. These combined strategies (Fullan, 1998) focussed the accountability processes on enhancing student achievement through program development and inclusive discussions with school and community stakeholders. These

strategies broadened the depth and breadth of persons involved supporting school programs and student learning. The nature of the relationship between the key strategies (DIP, SIP, public reports and consultations) reflected a focus on process, collective voices, policy adaptation and professionalism. Such cross-case findings lend support to Hargreaves' (1994) position that responsible decision-making in restructuring contexts is not centralized or decentralized but dispersed.

With shared responsibility and combined strategies in place, the cross-case findings help to illustrate that school response to EQAO requirements and recommendations was proactive in many ways: Teacher cooperation and professional dialogue, inherent in the accountability processes, supported school improvement by enabling professional learning to address misconceptions and problem-coping. EQAO recommendations to inform and engage parents, students and community stakeholders in assessment conversations were being realized. The use of performance tasks and a variety of assessment methods was common practice in both schools by the fourth year of the Grade 6 EQAO testing implementation. School-based accountability was established because reviews of school performance and ongoing improvement planning were established practice in both schools. The study findings clearly revealed how the aims and mechanisms of the internal and external system functioned in a compatible manner and supported a proactive school response. School response was mobilized by supportive professional networks, community agencies and cultural support groups. These networks extended the assessment conversation beyond the immediate school, its staff and principal, to include parents and community stakeholders. It can be said that ownership and commitment to school performance was more inclusive as a result.

The sixth and final discussion point unveils the influence of strong internal systems to promote practices indicative of assessment communities. The cross-case analysis revealed how structured opportunities for professional dialogue and discourse supported experimentation with new instructional materials and assessment methods. The findings provide evidence that collegial venues (staff and team meetings, mentoring, principal leadership and supervision) encouraged teachers to use new assessment practices and to seek understanding of the evaluative criteria pertinent to the provincial standards. The study findings also point to the presence of some formative assessment activities in both schools and disclose teacher engagement in learning about curricular expectations and interpreting student achievement with appropriate criteria. The study findings also suggest that teachers were using multiple methods to assess students' performance and were engaging in reflection on the assessment process. Assessment conversations were classroom-based but were also school-based as schools were actively engaged in ongoing reviews of school performance. These assessment activities facilitated school improvement and ensured that schools remained accountable on many levels. The extant internal accountability systems were found to be well positioned to support this continued attention.

In summary, the study findings reveal that strong internal accountability systems were functioning in both case study schools, and that these forums possessed all five components of the integrated model. The internal systems were found to be instrumental in enabling the dual aims of the EQAO agenda to be realized in each school in such a manner that local student needs and community priorities were addressed. Specifically, the internal accountability systems monitored how the school staff accounted to school

and community stakeholders for the instructional program and facilitated enhanced teachers' practices.

This chapter discussed the cross-case findings from this collective case study of internal accountability systems. How these findings inform the conclusions and contribute to theory and to practice is addressed in the final chapter.

CHAPTER 6

CONCLUSION

The purpose of this collective case study was to explore and characterize the nature of internal accountability systems, and to describe the mediating influence that these school-level forums have on teachers' assessment practices. This study anticipated systemic pressures and supports would illuminate how accountability was conceptualized in the province of Ontario and aimed to capture school response to the dual aims of the EQAO accountability agenda.

Many factors that support and hinder change in teachers' practice are reported in the literature (Borko, 1997; Chrispeels, 1997; Firestone et al., 1998; Grant et al., 1996; Peterson et al., 1996; Van der Vegt et al., 2001), particularly in contexts where large-scale testing is instituted to assess school performance. The findings of this collective case study acknowledges the need for timely arrival of resources and supports for curriculum renewal and the presence of anxiety related to some unintended negative effects of the testing and transparent reporting. This said, the intent of this research was to explore beyond these valid concerns and illuminate the structures and processes at play within school accountability contexts that enabled schools to carry out their responsibilities, solve problems and cope with the implementation of standards-based policies.

This focus was achieved by describing the components of the internal accountability systems and by characterizing teachers' assessment practices through analyses of interview data, field notes, and documentation collected from multiple

sources and presented in the preceding chapters. It was enabled by capturing teachers' collective (school) response to accountability mandates and recommendations.

This final chapter presents the conclusions based on the cross-case findings in an effort to answer the two research questions; *What is the nature of school internal accountability systems in a context of external accountability?* and *What are the mediating influences of internal accountability, if any, on teachers' assessment practices?* The study conclusions are viewed as tentative suppositions for consideration rather than definitive statements, since they are based on case studies of two elementary schools. What this exploratory study of internal accountability systems contributes to research is outlined first, then followed by several conclusions that highlight key study findings. The chapter then outlines the limitations of the study and explains the extent to which the findings may be applied. The chapter concludes by outlining the implications for research and for practice.

Contributions and Highlights

Much of the literature on internal accountability systems queries whether these school level forums can effectively meet the need for public accountability, and, at the same time, promote meaningful change in restructuring contexts with vigorous external accountability (Abelmann & Elmore, 1999; Earl & Torrance, 2000; Fuhrman, 1999; Hargreaves, 1993; Leithwood et al., 2002; Newman et al., 1997; Snider, 2001). These works speak to diverse forms of accountability frameworks, but refer to the emergence of a new form of accountability better situated to meet the demands of today's public education. The public demand for system accountability has stimulated a differentiation between reform as bureaucratic control versus reform as professional empowerment.

This study described the effectiveness of Ontario's professional practice approach to system-wide accountability and reports on its ability to support the kind of internal accountability that is associated with school improvement.

This qualitative study of internal accountability systems also contributes to theory by conceptualizing these school-based systems to include both structural and socio-cultural dimensions as presented in the integrated model. Empirical studies and reviews (Cohen, 1995; Darling-Hammond, 1990; Fuhrmann, 1995; Hargreaves, 1996; Knapp, 1997; Spillane, 1998) allude to the importance of considering both structural and cultural dimensions in understanding policy to practice connections. These works advise that systemic reform must address issues of professional learning in addition to structuring policy implementation. This is necessary so that policies are better understood and change processes are enabled by compatible strategies throughout the system.

This collective case study used multiple theoretical lenses to inform how policy and practice interact through the use of accountability mechanisms. It identifies the operant structures and processes that comprise these school-based forums. This study also provides insights on factors that influence decision-making and motivate teachers' commitment to change. The integrated model of internal accountability and the conceptual framework developed to guide this study builds upon the works of Abelman and Elmore (1999) and Newman et al. (1997). This positions the discussion of the findings from an inclusive theoretical perspective. This exploratory study was designed to address limitations inherent in earlier studies of system change that looked exclusively at either the dynamics of policy implementation or issues of school culture.

In using multiple lenses, policy to practice connections are disclosed and the limitations associated with a single theoretical perspective (Knapp, 1997) diminished.

Furthermore, this study situates its descriptions of teachers' assessment practices from what Brookhart (2004) explains as an intersection of three teaching functions: instruction, management, and assessment. The theory relevant to this study of teachers' assessment practices is from the field of measurement and evaluation as it relates to assessment, but also includes change theory as it relates to implementation of assessment policy and management of professional learning. This study uses this theoretical intersection to illuminate the link between standards policy and change in teachers' practices.

Newman et al. (1997) proposed that external accountability could play a supportive role in fostering the kind of internal accountability associated with quality instructional programs if external agents could offer concrete examples of high performance, promote assessment approaches and provide reliable measures of student performance (p. 61). Evidence from the cross-case analyses lends empirical support to this proposition.

The study findings provide evidence of the substantive contribution that the EQAO performance tests and accountability program made to the internal accountability systems in the case study schools. Specifically, the external measures of school performance provided information that contributed to ongoing evaluation (assessment) of the school instructional programs. The performance tests themselves modeled quality instructional tasks that demonstrated the curriculum expectations and reflected the importance of assessment as feedback in learning processes. The EQAO requirement

that SIPs become part of the annual review of school performance created an opportunity for schools and the public to engage in assessment conversations focussed on developing collective expectations and supporting continuous improvement. This is suggestive of a policy menu approach (Fullan, 1994) as the EQAO and district SIP guidelines set literacy and numeracy parameters, but not specific goals or activities. This responsibility was delegated to school-based professionals. These findings support the supposition that a balance between external oversight and school-level professional responsibility was extant. This balance, represented by the presence of compatibility instruments and strategies, suggests that policymakers at each system level respected professional values and owned responsible decision-making.

Abelmann and Elmore (1999) cautioned that strong external accountability systems would have little impact on internal accountability if schools did not develop collective expectations among the members of the school and its community. These researchers reported that schools often lacked understanding of what students can do and what influence of student, community and school characteristics have on student learning (Abelmann & Elmore, 1999, p. 42). This exploratory study sought evidence of such perspectives and unveiled what information was used to assess school performance in two Ontario schools. In contrast to other studies that showed diverse expectations, it was found that the schools in this study developed collective expectations, had concrete examples of student performances for assessing achievement, and took into consideration the many influences that impact student learning. These program assessment issues were being addressed by the operant accountability mechanisms (SIPs, EQAO tests, public reporting, professional dialogue, etc.) that guided and

monitored teachers' practices within each school. These processes reflect the substantive contribution of the EQAO program to enable the professional ownership for program review (evaluation) and justification. These processes appear to address what Leithwood et al. (2002) called local challenges to proactive change because they promoted assessment conversations that considered multiple measures of school performance, school and community characteristics, and program goals in a manner designed to extend ownership for students' learning while respecting teachers' responsibility for quality instructional programs.

This research supports the proposition (Newman et al., 1997) that strong internal accountability systems possess common structural components, including information on student performance, a standard of achievement, persons who judge school performance, and consequences when school performance doesn't meet expectations. The cross-case findings also support the proposition that these internal systems possess aims (collective expectations) that consider issues of student potential and environmental factors, but that rely on interactive processes (mechanisms) to guide teachers' practice. This study therefore proposes that consideration be given to a conception of internal accountability that employs both structural and socio-cultural dimensions. The integrated model, outlined in the conceptual framework and used to analyze the case study data, guided the identification of those elements that comprised the five components of the internal systems.

The study findings also support the proposition that internal systems possess common components but may possess different configurations. The findings suggest that variations in some of the elements (Abelmann & Elmore, 1999) are a function of diverse

school contexts and community priorities. In a sense, this variation reveals the flexibility of the Ontario accountability system to respect school and community differences, while promoting professional responsibility to address student achievement. The subtle differences in school management did not detract from the presence or the effectiveness of the five components framing the infrastructure of the internal system.

This study provided a unique opportunity to investigate the nature of internal accountability systems nested within a state-level accountability environment that favored a professional practice approach. The findings of this study endorse the position (Earl & LeMahieu, 1997; Earl & Torrance, 2000) that large-scale tests are able to support the dual aims of *assessment as learning* and *assessment as accountability*. The findings illustrate that assessment conversations between teachers, schools, parents and community groups are more inclusive at the school level than previously reported in early studies (Earl & Torrance, 2000; Snider, 2001) of Ontario's accountability initiative. Furthermore, this study characterized teachers' assessment practices and found school assessment practice possessed many characteristics associated with an assessment community (Hall, Harding et al. 2002). These findings reveal that institutional development (organizational effectiveness and professional learning) was facilitated by internal accountability systems, and hint that the participating schools appeared to be operating at different transitional levels in terms of full implementation of standards-based assessment. Using categories reported in Hibbard's (1996) frontier model, both pioneers and settlers are present and working to ensure that that varied assessment purposes and multiple methods were part of teachers' practices. With performance based testing at the Grade 6 level in its fourth year of implementation, the cross-case findings

pointed to enhanced assessment practices and school accountability as being present in both schools.

The scope of this exploratory study cannot ascertain if these practices are institutionalized to the point of sustainability. However, several criteria associated with sustaining large-scale reforms (Fullan, 2000) are reported in the study findings. These include the standards-based EQAO testing program, the provincial achievement standards and curriculum expectations, the compatible accountability instruments and strategies, the networking between professional groups, the availability of quality resources, and extended ownership for school improvement.

In considering the influence of strong internal accountability systems on teachers' practices, this study examined the characteristics of classroom assessment and teachers' collective assessment practices. In first describing the purpose, criteria, methods, and feedback of classroom assessment and then identifying the mechanisms that promoted teacher learning, the cross case findings revealed the collective dynamics indicative of emerging assessment communities. The cross-case findings link such processes to the policy structures and interactive strategies of the internal accountability systems.

Ontario's current accountability agenda was designed to address two aims that at first glance may appear to be quite diverse. The first aim was to make schools and their teachers accountable for the instructional program that they offered through the use of an accountability agenda tied to large-scale provincial testing. The second aim was to support teacher learning and enhanced teaching through the design of these performance-based tests. The cross-case findings support the conclusion that these two

aims were being realized in both schools. The findings also support the position that strong internal accountability systems played a pivotal role in the process. The highlights of this collective case study are now organized in the following seven conclusions.

The first conclusion is that internal accountability systems are school forums comprised of five components parts. The parts include: aims, information on school performance, a standard of achievement, persons who judge school performance, and mechanisms that guide and monitor teacher practice. These systems include both structural and socio-cultural dimensions that define organizational responsibilities and procedures that foster policy to practice connections.

The second conclusion is that strong internal accountability systems are possible within active external accountability systems intending system change. This conclusion supports the proposition that external accountability policies and mechanisms can guide and support school-based accountability. Furthermore, these systems appear to possess common components, specifically compatible external and internal strategies associated with Ontario's accountability program. A degree of variation in the elements of some components is possible without detracting from the school-based accountability or reducing the compatibility between the external and internal strategies. This finding suggests that internal accountability systems have some flexibility to adapt external requirements and recommendations to address local priorities and student needs.

The findings also support a third conclusion: strong internal accountability can be found within external accountability frameworks employing large-scale performance testing programs when, as in this study, the provincial accountability framework is based upon standards principles and values professionalism. This supports the proposition that

accountability testing can possess dual aims (accounting and learning). This appears possible because the underlying beliefs and assumptions that formed the conception of accountability in Ontario are based on democratic principles and moral sources of authority, a perspective necessary to enable responsible decision-making.

The fourth conclusion arising from this collective case study is that strong internal accountability systems can mediate classroom assessment practice and support the development of collegial assessment practices. This finding supports the proposition that internal accountability systems may be effective forums for encouraging collective (teachers') practice because in both schools such systems prompted collegial dialogue and professional learning.

A fifth conclusion supported by the findings is that responsible decision-making extended the participation and commitment of stakeholders to school improvement by including teachers, administrators, parents, and community leaders in program reviews and goal setting. The SIP processes and the School Reports of school performance encouraged principals and teachers to explain and justify school programming and consult with other stakeholders. These consultations built understanding of collective expectations and rallied support. In doing so, ownership and commitment to student learning and school performance was more inclusive. The responsibility and involvement of actors at the state, district and local levels were differentiated, resulting in a balance between centralized and decentralized control.

A sixth conclusion is that compatible (coherent, integrated) strategies facilitate policy to practice connections across the provincial, district and school level of the education system. The findings reveal that accountability strategies such as instruments

and processes were process-oriented and designed to engage the voice of school and community stakeholders. Furthermore, these strategies respected professionalism because instruments like the SIP provided structure but not specific actions or outcomes ensuring a degree of adaptability to local needs and priorities. The use of compatible external and internal strategies does not seem to preclude a degree of horizontal variation among schools. The integrated and process-oriented aspects of the DIP, SIP, school reviews, and public reports required that explanations and justification of instructional decisions be shared with all school stakeholders. These requirements also broaden involvement in accountability processes. This study's findings disclose that such strategies are able to broaden the depth and breadth of persons committed to school improvement.

A final conclusion based on the study findings is that two environmental factors impacted the internal accountability systems positively. These included the EQAO testing as a policy lever and school leadership facilitating mobilizing supportive actors. The EQAO tests provided an external standardized measure of school performance that clarified curriculum expectations and modeled performance-based tasks. In both case studies, the principal was also pivotal to facilitating the internal accountability system processes and procedures. Although not a focus of this study, district level support to principals facilitated professional networking and enabled school-community collaboration.

These seven conclusions elucidate the role that internal accountability systems played in supporting school accountability for quality instructional programs and for enhancing teachers' assessment practice. These conclusions highlight how such school-

based forums are a means to school improvement. In light of the study findings, it would be prudent for educators and policy makers to investigate accountability policy frameworks and strategies that are conducive to the development of strong internal accountability systems. The fact that such systems have been found to augment *accountability as assessment* and *accountability as learning* advances emerging theoretical propositions about the role of professional models of accountability in systemic change and highlights the role large-scale performance testing can play in school improvement.

Limitations

In reflecting on the limitations of this work, the project is first bounded by the scope and scale of the research design. The intent of this study was to explore the nature of internal accountability systems and to describe school level response within two elementary school contexts during a period of standards-based policy implementation.

This study is but a snapshot of an ongoing implementation of standards policy in the fourth year of a large-scale performance-based test at the Grade 6 level. The study examined the use of these tests as a policy lever to ensure elementary schools were accountable for their instructional programs and engaged in enhancing teachers' practice. The study explored how the EQAO accountability program tied these tests to other instruments and strategies that facilitated the development of internal (school-based) accountability systems in elementary schools. As it does not address the nature of accountability at the high school level, these findings may not be applicable where accountability testing is tied to students' certification.

Change takes time to initiate, implement and sustain; and many factors impact on the process. This study examined the status and the nature of internal accountability systems at one point in time. It analysed school level response to recommendations about how schools do account to their community stakeholders. While describing classroom assessment practices and capturing the degree to which characteristics associated with assessment communities were established, this study cannot claim that these practices were sustainable over time or specifically determine how much change can be attributed to the EQAO's accountability agenda. These assertions are limited by the scope of this study and rely on participants' perspectives and reports.

The study does capture the structure and dynamics of the internal system components. It also illuminates the nature of internal accountability in a state-level framework valuing professional practice. However, the study cannot claim the same elements would be found in different accountability frameworks but suggests that the same components can be anticipated.

In addition to the limitation of scope, this study is limited by scale, as two cases do not provide a comprehensive picture of the possible configurations of internal accountability. Nor does this study propose that all schools will be successful in developing characteristics of inclusive accountability and collegial assessment practice by promoting active accountability forums as described here in. However, the experiences of these two schools can provide educators and interested others with some understanding of school contexts and conditions that facilitate proactive school-level responses to external demands for accountability and enhanced teaching practices.

A third aspect of this collective case study that may be considered a limitation was the choice of a district that possessed schools with high capacity. This selection was purposive in order to limit the variables that might impact on the presence of internal accountability systems to those associated with policy implementation and organizational change. This design choice enabled the analysis to focus on the external accountability policy instruments and mandates and the school organizational dynamics such as procedures, routines and patterns of behaviour, in order to determine what components of the integrated model were present and to characterize the configuration of these elements. This is not to say that some degree of variability in individual teacher capacity was not present, but that, as a collective entity, the skills of the teachers, principals and district support staff were characteristic of high capacity.

A fourth potential limitation was the presence of researcher bias. Since the research is interpretive in nature, how events were perceived and data were analyzed are considered from the perspective of an experienced educator, with knowledge of schools and educational institutions developed over an extensive professional career. The study's design attempts to deal with possible bias by using triangulation based upon member-checking, cross referencing among divergent data sources, and thick description incorporating the words of the participants. Reflective dialogue with an advisor and continued readings were also used. In a sense, the professional acquaintance of the researcher with schools and the EQAO tests is felt to have contributed to an appreciation of the complexity of school organizations and the myriad of professional interactions associated with school accountability and teachers' practices.

Implications for Research

Reports in the literature vetting new forms of accountability are emerging. Implications from this study point to the effectiveness of a professional practice model of accountability that employs large-scale assessments to promote school accountability and to facilitate change in teachers' practices. Ontario's model of accountability is one informed by empirical research on innovation and change as well as the field of policy implementation. It takes into account knowledge of the research on professional learning and organizational development. The dual aims of making schools accountable for their instructional programs and for promoting change in teachers' practice were structured into the policies, instruments and strategies of the accountability framework designed by the EQAO, the provincial organization responsible for assuring quality in the Ontario education system.

This case study research specifically explored the impact of the Grade 6 accountability testing on two elementary schools. The study findings speak to the existence of school internal accountability systems and to the reported changes in teachers' assessment practice. These findings were made as the accountability testing was in its fourth year of implementation. The professional practice model of accountability employed in Ontario was found to inform teachers' practice and to require schools to explain and justify their instructional programs based on the study findings. Internal accountability systems played a role in this achievement.

Qualitative studies examining the nature of school internal accountability systems in such contexts are limited, especially in nested state-district contexts promoting responsible decision-making and the involvement of school and community

stakeholders. The findings of this exploratory study revealed strong internal accountability systems were operating in both schools. The findings provide evidence that strong internal accountability systems are associated with processes that support teacher learning and school improvement as well as with structures that ensure public accountability.

To better examine the configurations of internal accountability systems and to explore the factors that impact such systems in different districts or school communities warrants further study in Ontario schools. Qualitative designs would inform on the elements that vary between different configurations, and develop understanding of the full range of strategies and procedures. Research that uses comparative methodology and purposively selects high versus low capacity contexts may elucidate the significance of specific contextual conditions to the establishment of strong internal accountability.

Large-scale studies of school-based accountability systems are required before generalizations about commonalities in the system components are possible, and whether different configurations are stable. Questions remain about the nature of internal systems and *how they may change over time*. This suggests implications for longitudinal studies that examine highly developed versus weak internal accountability systems. Are there patterns associated with transitional stages? Are there plateaus in the process of policy implementation and why? What type of leadership is beneficial within each stage of the process? Will patterns in mechanisms evolve over time as goals are met and new ones identified? What patterns in configuration are representative of emergent systems, stable systems and established systems? Do these configurations always vary or are they characteristic of different school communities? Considering the mediating influence of

internal accountability systems on teachers' practices, these areas are worthy of continued investigation.

Another implication for research relates to the role of intermediate level accountability systems (district level accountability). The intermediate and supportive role of the school district in the accountability framework has been explored in this study, but in a modest way. Much needs to be understood about the management of district policy and of district support structures. As school internal accountability systems are examined, similarities and differences in district accountability systems would appear worthy of exploration as well. How do different districts perceive their accountability function, and what are the aims, priorities and strategies of their action plans? How do different districts support program review and strategic planning? What kinds of strategies are employed, and are intermediate policies and instruments aligned with provincial accountability strategies? What similarities and differences exist between district structures and community cultures? Further investigations of schools within a common district would provide insights into the nature of district support; while studies comparing district contexts nested within the same state-level environment could delineate what district level actions enable or hinder the development of school-based accountability.

Another arising issue suggests that much needs to be learned about the unintended negative effects of large-scale testing in such transparent systems when changes in government occur and shift in political will impacts on the original aims of the restructuring agenda. One query arising from the study was the impact of accountability testing when individual student performance on high school literacy and

numeracy tests is used for graduation purposes. Will students and schools from disadvantaged communities be better served in the long run? Will funding and resources be channelled to those schools most in need of additional supports?

Related to student achievement and certification is the issue of parental involvement in the development of individual student learning plans. Although parental input at the school advisory level was well defined, less is known about the nature of parental involvement at the classroom level, or what parents' perceived to be their role in supporting student learning. Teachers reported asking for parent support for homework and ongoing review to support the expectations seen as somewhat ambitious. Little is known to date of what parents understand about the school accountability processes and the role of formative assessment in school improvement. Parent perspectives of assessment and their participation, if any, in planning their children's individual educational plans are little understood in Ontario's context. In an accountability framework that extends involvement and aims to bring in community and parent participation, a study that examines this dynamic is warranted, likely at a point in time when the transition to an assessment community and implementation of standards-based principles have been established. Is parent involvement symbolic or is it critical to the development of assessment communities? Is the definition of an assessment community fully developed at this point in time? What type of activities are parents willing and able to commit to in supporting their children's schooling.

Governance issues were not addressed within this study, but issues related to funding and resources did emerge. Changes in school and district funding formulas did impact the availability of resources, and the presence of special provincial grants did

represent external guidance on where funds should be spent and what were important provincial priorities. Professional re-certification was another area of discourse falling within this domain. The impact of such issues on the stability of the Ontario accountability framework are yet unknown and are worthy of closer examination.

Implications for Practice

This collective case study of two elementary schools involved in a system-wide restructuring agenda contributes to knowledge about how individual schools responded to an agenda of accountability as assessment and as learning. Several implications for teachers' practices are inherent in the findings because a professional practice approach to educational accountability was being employed.

An effective accountability framework is a collaborative enterprise where policy to practice connections are supported by common aims and compatibles strategies and processes. Such an endeavour relies upon a common set of underlying assumptions and beliefs about the purpose of schools and what they are accountable for and to whom. In Ontario, the accountability agenda was designed in response to a province wide consultation among all stakeholders, and the current approach is premised on a model of professionalism rooted in democratic principles and moral sources of authority.

This professional practice model means schools, the principals and the teachers, must provide an account of what they do and why they do it to parents, the community and other professionals within the system. This accounting requires explanation and justification of educational programs and instructional practices. The cross-case findings revealed that the case study schools responded to the accountability framework by

developing strong internal accountability systems and that these systems mediated teachers' collective assessment practices. Five implications for teachers' practice are indicated in the findings. These are described in the following paragraphs.

First, internal accountability systems are important school-level forums because they possess internal and external mechanisms for examining program effectiveness and for stimulating improved teachers' practices. These systems possess specific component parts that contribute to a comprehensive assessment of school performance that predicates ongoing school improvement planning. Such systems possess a variety of mechanisms where teachers' practice is guided and monitored; with some externally mandated while others internally generated. These combined strategies revealed how a balance between external oversight and local support could be achieved while ensuring that schools remained accountable. These systems are informed by external policy instruments such as large-scale testing programs, but are enabled by school leadership and professional values.

Strong internal accountability systems do appear to possess common components and similar elements, but this does not mean that all features are identical or that they should be. Internal systems are site-based forums and likely reflect characteristics that are specific to a school context, its community priorities and identified student needs. School professionals are key agents in shaping the nature of this lateral accountability (Fullan, 2000) because school aims and goals are constructed in the collegial dialogue and in the assessment conversations among school and community stakeholders. The cross-case findings showed that principals, teachers, parents and community leaders participated in the account-making processes, but the

analysis also implies that teachers were not fully cognizant of internal accountability as a functioning, school-based forum that allowed staff to make sense of multiple innovations and take ownership for how changes were carried out. Conceptualizing internal accountability and promoting professional ownership of this school-based forum is one implication for teachers' practice arising from this study.

Second, the role of school leaders in managing the internal accountability system is another implication for practice. School leaders must value the cooperative and collaborative processes needed to develop collegial assessment. Their actions need to support professional learning situated in teacher dialogue and discourse. Fundamental to this school improvement is transformational leadership that establishes a collective vision, mobilizes supportive resources and provides individualized support to teachers. School leaders are key to effective school-based professional learning (e.g., best practices) and should consider modeling the type of formative assessment (e.g., program review, assessment conversations) on a professional level that teachers should be using to scaffold students' learning in the classroom. Teachers need to feel comfortable experimenting with multiple assessment methods and have opportunities to collegial discuss evaluative criteria in the course of planning and teaching. Principals, as managers and leaders, must pace organizational changes so that commitment to school improvement is meaningful and sustainable.

Another implication for practice relates to the depth and breadth of persons involved in the assessment conversations. Mobilizing resources and establishing school-community networks supports school-based accountability because these environmental conditions provide pressure and support for professional development. Extending

participation requires sensitivity to community perspectives and awareness of the capabilities of the participants. It also means that principals and school leaders must be confident in their understanding of accountability processes and knowledgeable about the principles of standards-based policy. This study may be helpful in this regard. Connecting with district, school and community stakeholders extends ownership and accountability for students' learning and school leaders should consider being proactive in these areas.

A fourth implication for practice from this work refers to the use of internal accountability systems to promote the development of assessment communities. Standards-based assessment is a collective exercise. This means that principals must promote the use of diverse assessment methods and expect ongoing measures of student progress by facilitating teaching partnerships (mentors). Assessment communities build assessment literacy and foster collective expectations through discussion of criteria and review of exemplars. School improvement benefits from moderated assessment as much as from good instruction. A strong internal accountability system helps to maintain focus on achievement in light of, not in lieu of, other student needs and community priorities. Principals and teachers require opportunities to develop collective understandings of what moderated assessment means and how it can benefit students' learning. Internal accountability systems are forums to address these local challenges to school improvement.

The study findings disclosed the supportive role of the district in gathering assessment data and allocating resources. School improvement benefits from technical support and professional guidance from district experts. The findings identified that

district SIP in-service, assessment and grading handbooks and school audit guidelines informed school-based decision-making. These district initiatives helped make school review processes meaningful exchanges of ideas and problem-solving. Instead of being passive recipients of school performance data, district support for school leaders and teachers can trigger professional confidence to use these assessment conversations as opportunities to showcase what they do in their classrooms. Such a goal requires school-based professional development that is ongoing and pertinent to the teachers. Networking between schools, professional groups and community agencies has obvious advantages. These consultative venues facilitate problem-solving and teachers' learning.

Parental involvement in assessment conversations is important for several reasons. Among them is the opportunity to explain the role of assessment within instruction, to clarify the function of an ambitious achievement standard, and to identify how school-community indicators impact student achievement. These interactions require explanations and justification of instructional programs - a high level of accountability and are potential avenues to bring students and parents into planning next steps and motivating students by demystifying assessment and grading.

As the fifth and final implication for practice, teachers are reminded that large-scale testing must be viewed as one source of information on school performance. Its contribution is unique because it provides an external measure of program effectiveness. Results from these tests provide a system-wide perspective that can contribute to program evaluations because it factors equity and excellence based on common curriculum expectations into the review process. However, schools must consider these testing results in light of other sources of information impacting student achievement so

that evaluating the effectiveness of the instructional program is based on determinations of *value added* rather than arbitrary numerical comparisons. The importance of extensive data gathering (EQUIP, aptitude tests, etc.) is relevant because appropriate goals and suitable improvement plans need to be crafted. Regular review of school performance is also necessary so that the effectiveness of improvement plans can be measured and ongoing problem-solving guided.

Strong internal accountability systems are associated with schools and districts that possess high capacity and this was a fixed variable in this study's design. This characteristic is an important requirement for a professional model of accountability based on democratic principles and moral sources of authority. Responsible decision-making at each level of a restructuring education system requires professional expertise to function collaboratively. The study findings support the effectiveness of such an approach to promote change in teachers' practice with the proviso that professional learning is an integral part of school improvement activities. The findings confirm that in the case of the two schools studied strong internal accountability systems are associated with high capacity. Since these systems benefit schools in many ways, establishing internal accountability systems is an important professional responsibility.

References

- Abelmann, C., & Elmore, R. (1999). *When accountability knocks, will anyone answer?* Philadelphia: Consortium for Policy Research in Education. (ERIC Document Reproduction Service No. ED428463).
- Adams, J., & Kirst, M. (1999). New demands and concepts for educational accountability: Striving for results in an era of excellence. In J. Murphy & K.S. Louis (Eds.), *Handbook of research on educational administration*. San Francisco: Jossey-Bass.
- Airasian, P.W. (2000). *Assessment in the classroom: A concise approach*. Boston: McGraw-Hill.
- Arter, J. (1999). Teaching about performance assessment. *Educational Measurement: Issues and Practices*, Winter 1999.
- Barnett, K., & McCormick, J. (2003). Vision, relationships and teacher motivation: a case study. *Journal of Educational Administration*, 41(1), 55-73.
- Biggs, J. (1995). Assessing for learning: Some dimensions underlying new approaches to educational assessment. *Alberta Journal of Educational Research*, 61 (1), 1-17.
- Black, P. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, 21, 49-97.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5 (1), 7-74.
- Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment:

Successes, stumbling blocks, and implications for professional development.

Teaching and Teacher Education, 13(3), 259-278.

Brookhart, S.M. (2004). Classroom assessment: Tensions and intersection in theory and practice. *Teachers College Record*, 106(3), 429-458.

Chrispeels, J. (1997). Educational policy implementation in a shifting policy climate: The california experience. *American Educational Research Journal*, 34 (3), 453-481.

Cibulka, J. (1991). Educational accountability reforms: performance information and political power. In S. Fuhrman & B. Malen (Eds.), *The Politics of curriculum and testing*. New York: Falmer Press.

Cizek, G., Fitzgerald, S., & Rachor, R. (1996). Teachers' assessment practices: Preparation, isolation and the kitchen sink. *Educational Assessment*, 3 (2), 159-179.

Claussen, K. (2001). *The meaning and implementation of curriculum integration in the middle school years: Ministry of Education reforms and current practice in the Ontario school system*. Unpublished doctoral dissertation, University of Ottawa, Ontario.

Cohen, D. (1995). What is the system in systemic reform? *Educational Researcher*, 24 (9), 11-17.

Cohen, D., & Ball, D. (1990). Policy and practice: An overview. *Educational Evaluation and Policy Analysis*, 12(3), 347-353.

Corcoran, T., & Goertz, M. (1995). Instructional capacity and high performing schools. *Educational Researcher*, 24 (9), 27-31.

- Costa, A., & Kallick, B. (Eds.). (1995). *Assessment in the learning organization: shifting the paradigm*. Virginia: ASCD.
- Cousins, J.B. (1996). Understanding organizational learning for educational leadership and school reform. In K. Leithwood et al. (Eds.), *International handbook of educational leadership and administration*, (pp.589-652). Netherlands: Kluwer Press.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 438-481.
- Cross, C., & Joftus, S. (1997). Are academic standards a threat or opportunity? *NASSP Bulletin*, September 1997.
- Daniel, L.G., & King, D.A. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary teachers. *Journal of Educational Research*, 91(6), 331-344.
- Darling-Hammond, L. (1990). Instructional policy into practice: The power of the bottom over the top. *Educational Evaluation and Policy Analysis*, 12 (3), 233-241.
- Darling-Hammond, L., & Ascher, C. (1991). *Creating accountability in big city schools*. New York: National Center for Restructuring Education, Schools and Teaching.
- Datnow, A., & Castellano, M.E. (2001). Managing and guiding school reform: Leadership in success for all schools. *Educational Administration Quarterly*, 37(2), 219-249.
- Earl, L. (1995). Assessment and accountability in education in ontario. *Canadian Journal of Education*, 20, (1), 45-55.

- Earl, L., & Cousins, J.B. (1996). *Classroom assessment: Changing the face and facing the change*. Commissioned by the Ontario Public Schools Teacher Federation.
- Earl, L., & LeMahieu, P. (1997). Rethinking assessment and accountability. In A. Hargeaves, (Ed.). *Rethinking educational change with heart and mind*. Alexandria: ASCD.
- Earl, L., & Torrance, N. (2000). Embedding accountability and improvement into large-scale assessment: What difference does it make? *Peabody Journal of Education*, 75(4), 114-141.
- Elmore, R.F. (1990). Backwards mapping: Implementation research and policy decisions. *Political Science Quarterly*, 94 (4), 601-616.
- Elmore, R.F. (1995). Structural Reform and Educational Practice. *Educational Researcher*, 24 (9), 223-26.
- EQAO. (1997). Provincial Report on Achievement. Toronto, ON: Author
- Firestone, W.A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in maine and maryland. *Educational Evaluation and Policy Analysis*, 20 (2), 95-113.
- Fuhrman, S. (1995). Introduction: Recent research on education reform. *Educational Researcher*, 24 (9), 4-5.
- Fuhrman, S. (1999). *The new accountability*. Philadelphia: University of Pennsylvania. (CPRE Research Report No. RB-27)
- Fuhrman, S., & O'Day, J. (1996). *Rewards and Reform*. San Francisco : Jossey-Bass Publishers.

- Fullan, M. (1993). *Change Forces: Probing the Depths of Educational Reform*. Philadelphia: Falmer Press.
- Fullan, M. (1998). Education reform: Are we on the right track? *Education Canada*, 38(3), 4-7.
- Fullan, M. (2003). *The moral imperative of school leadership*. New York: Corwin Press.
- Fullan, M. (1991). *The new meaning of educational change*. Toronto: OISE Press.
- Geijsel, F., Slegers, P., Leithwood, K., & Jantzi, D. (2003). Transformational leadership effects on teachers' commitment and effort toward school reform. *Journal of Educational Administration*, 41(3), 228-256.
- Grant, A., Peterson, P., & Shojgreen-Downer, I. (1996). Learning to teach mathematics in the context of systemic reform. *American Educational Research Journal*, 33(2), 509-541.
- Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. Washington, DC: Falmer Press.
- Green, J. (1998). Authentic assessment: Constructing the way forward for all students. *Education Canada*, 38(3), 8-12.
- Greenfield, W. (2004). Moral leadership in schools. *Journal of Educational Administration*, 42(2), 174-196.
- Gurney, P.J., & Andrews, B.W. (2002). Restructuring faculties of education: State of the nation. *International Electronic Journal for Leadership in Learning*. Retrieved July 12, 2004, from <http://www.ucalgary.ca/~iejll>.
- Haertel, H. (1999). Validity arguments for h-stakes testing: In search of evidence. *Educational Measurements: Issues and Practices*, Winter 99, 5-9.

- Hall, K., & Harding, A. (2002). Level descriptions and teacher assessment in England: Towards a community of assessment practice. *Educational Researcher*, 44(1), 1-15.
- Hargreaves, A. (1992). Cultures of teaching: a focus on change. In A. Hargreaves & M. Fullan (Eds.), *Understanding teacher development*. NY: Teachers College Press.
- Hargreaves, A. (1994). Restructuring restructuring: Post modernity and the prospects of educational change. *Journal of Educational Policy*, 9 (1), 47-65.
- Hargreaves, A. (1996). Transforming knowledge: Blurring the boundaries between research, policy, and practice. *Educational Evaluation and Policy Analysis*, 18(2), 105-122.
- Hargreaves, A., & Fullan, F.G. (1992). *Understanding teacher development*. New York: Teachers College Press.
- Harlen, W. (1994). *Enhancing quality in assessment*. London: Paul Chapman.
- Harris, A. (2000). What works in school improvement? Lessons from the field and future directions. *Educational Researcher*, 42(1), 1-11.
- Hatch, T. (1998). The differences in theory that matter in the practice of school improvement. *Educational Evaluation and Policy Analysis*, 35 (1), 3-31.
- Herman, J. (1997). Assessing new assessments: How do they measure up? *Theory into Practice*, 36 (4), 196-204.
- Hibbard, M. (1996). *A teacher's guide to performance-based learning and assessment*. Alexandria: ASCD.
- Jantzi, D., & Leithwood, K. (1996). Towards an explanation of variation in teachers' perceptions of transformational school leadership. *Educational Administration Quarterly*, 32(4), 512-538.

- Johnson, B., & Christensen, L. (2004). *Educational research: quantitative, qualitative, and mixed approaches*. Boston, MA: Pearson Education Inc.
- Joint Committee on Standards for Educational Evaluation [JCSSEE]. (2003). *The student evaluation standards: How to improve evaluations of students*. Thousand Oaks, CA: Corwin Press.
- Keene, E.A. (1995). The Implementation and impact of a state-wide performance assessment policy. *Dissertation Abstracts International*, (UMI No. 9525489)
- Knapp, M. (1997). Between systemic reform and the mathematics and science classroom: The dynamics of innovation, implementation and professional learning. *Review of Educational Research*, 67 (2), 227-266.
- Kruse, S.D. (2001). Creating communities of reform: Continuous improvement planning teams. *Journal of Educational Administration*, 39(4), 359-383.
- Kulm, G. (1994). *Mathematics assessment: What works in the classroom*. San Francisco: Jossey-Bass.
- Kulm, G., & Stuessy, C. (1991). Assessment in science and mathematics educational reform. In G. Kulm & S. Malcom (Eds.). *Science assessment in the service for reform*. Washington: AAAS.
- Lapointe, M.A. (2000). Is there a role for parents in school site accountability? A look at two schools of choice. *Dissertation Abstracts International*, 61(08), 3012A. (UMI No. 9981887)
- Leithwood, K. & Earl, L. (2000). Educational accountability affects, international perspective. *Peabody Journal of Education*, 74(4), 1-18.

- Leithwood, K., Edge, K., & Jantzi, D. (1999). *Educational accountability: The state of the art*. Gütersloh: Bertelsman Foundation Publishers.
- Leithwood, K., & Jantzi, D. (2000). The effects of transformational leadership on organizational conditions and student engagement with school. *Journal of Educational Administration*, 38(2), 112-129.
- Linn, R. (1993). Educational assessment: Expanded expectations and challenges. *Educational Evaluation and Policy Analysis*, 15 (1), 1-16.
- Linn, R. (1998). *Assessments and accountability*. Retrieved from <http://cresst96.cse.ucla.edu/CRESST/REPORTS/TECH490>.
- Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Linn, R., & Gronlund, N. (1995). *Measurement and Assessment in Teaching*. New Jersey: Prentice-Hall.
- Little, J. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis*, 15(2), 129-151.
- Lock, C. (2001). *The influence of a large-scale assessment program on classroom practices*. Unpublished doctoral dissertation, Queen's University, Ontario.
- Louis, K.S., & Marks, H.M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 10 (6), 532-575.
- Mawhinney, H. (1998). *Patterns of social control in assessment practices in canadian frameworks for accountability in education*. Paper presented at the meeting of the A.E.R.A., Atlanta, GA.

- Mawhinney, H. (1999). Rumbings in the cracks in conventional conceptions of school organizations. *Educational Administration Quarterly*, 35(4), 573-593.
- McEwen, N. (1995). Accountability in education in canada. *Canadian Journal of Education*, 20(1), 1-17.
- McLaughlin, M.W. (1987). Learning from experience: lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9 (2), 171-178.
- Merriam, B. (2002). *Qualitative research in practice*. San Francisco: Jossey-Bass.
- Meyer, H. (2002). From "loose coupling" to "tight management"? Making sense of the changing landscape in management and organization theory. *Journal of Educational Administration*, 40 (6), 515-520.
- Miles, A.M., & Hubermann, M.B. (1994). *Qualitative data analysis: An expanded sourcebook*. Newbury Park, CA: Sage.
- Mintrop, H., MacLellan, A.M., & Quintero, M.F. (2001). School improvement plans in schools on probation: A comparative content analysis across three accountability systems. *Educational Administration Quarterly*, 37(2), 197-218.
- Mitchell, K. (1997). What happens when school reform and accountability testing meet? *Theory into Practice*, 36 (4), 62-65.
- Monson, M. (1998). Performance assessment: Profiles in dilemma management. Harvard University. *Dissertation Abstracts International*, 59 (09), 3307A. (UMI No. 9908880)
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22(2), 155-175.

- Newman, F., King, M.F., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review*, 67 (1), 41-74.
- Ontario Ministry of Education (1994). *For the love of learning*. Toronto, ON: Author.
- O'Connor, K. (1999). *The mindful school: How to grade for learning*. Arlington Heights, IL: Skylight Professional Development.
- Peterson, P.L. (1990). The california study of elemental mathematics. *Educational Evaluation and Policy Analysis*, 12(3), 257-261.
- Peterson, P.L., McCarthy, S.J., & Elmore, R.F. (1996). Learning from school restructuring. *American Educational Research Journal*, 33(1), 119-153.
- Popham, W.J. (1999) *Classroom assessment: What teachers need to know*. Toronto, ON: Allyn & Bacon.
- Prestine, N.A., & McGreal, T.L. (1997). Fragile changes, sturdy lives: Implementing authentic assessment in schools. *Educational Administration Quarterly*, 33(3), 371-400.
- Quinn, D. (2002). The impact of principal leadership on instructional practice and student engagement. *Journal of Educational Administration*, 40(5), 447-467.
- Resnick, L., & Resnick, D. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. Gifford & M.C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction*. Boston: Kluwer Press.
- Rothman, R. (1995). *Measuring up: standards, assessments and school reform*. San Francisco: Jossey Bass.

- Ruohotie, P. (1996). Professional growth and development. In K. Leithwood et al. (Eds.), *International handbook of educational leadership and administration* (pp.419-445). Netherlands: Kluwer Academic Publishers.
- Ryan, A. (1997). Professional obligation: A dimension of how teachers evaluate their students. *Journal of Curriculum and Supervision*, 12 (2), 118-134.
- Sadler, R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education*, 13 (2), 192-208.
- Sergiovanni, T. (1996). *Leadership for the schoolhouse: How is it different? how is it important?* San Francisco: Jossey-Bass.
- Schafer, W.D. (1991). Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice*, Spring 1991, 3-6.
- Shepard, L.A. (1995). Using assessment to improve learning. *Educational Leadership*, 52 (2), 38-43.
- Shepard, L.A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, 16 (2), 5-13.
- Smith, M., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233-268). Bristol, PA: Falmer Press.
- Snider, H. (2001). *Reform assessment in ontario schools. Dissertation Abstracts International*. (DMI No. AAT 3004718)
- Spillane, J. P. (1998). State policy and the non-monolithic nature of the local school district: Organizational and professional considerations. *American Educational Research Journal*, 35 (1), 33-63.

- Spillane, J., & Thompson, C. (1997). Reconstructing conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. *Educational Evaluation and Policy Analysis, 19* (2), 185-203.
- Stake, R. (2000). Case studies. In N. Denzin & Y. Lincoln (Eds.), *The handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Stiggins, R.J., & Bridgeford, N.J. (1986) Classroom assessment: A key to effective education. *Educational Measurement: Issues and Practice*, Summer 1986, 5-17.
- Stiggins, R.J., & Conklin, N.F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. New York: SUNY Press.
- Van der Vegt, R.V., Smyth, L.F., & Vandeberghe, R. (2001). Implementing educational policy at the school level: Organizational dynamics and teacher concerns. *Journal of Educational Administration, 39*(1), 8-23.
- Wilson, (1990). Classroom processes in evaluating student achievement. *The Alberta Journal of Educational Research, 36*(1), 4-17.
- Yinn (1994). *Case study research*. Thousand Oaks, CA: Sage.
- Youngs, P., & King, M.B. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly, 38*(5), 643-670.
- Yu, H., Leithwood, K., & Jantzi, D. (2002). The effects of transformational leadership on teachers' commitment to change in hong kong. *Journal of Educational Administration, 40*(4), 368-389.

Appendix A

INFORMED CONSENT
(Teacher participant)

Principal Investigator: Adell Hay, Faculty of Education,
University of Ottawa
Address: Rm. 369 Lamoureux Hall, Faculty of Education
Telephone: (613) 599-XXXX (h)
Email: xxxxxxxxxxxx

I am currently a Ph.D. candidate at the University of Ottawa working under the supervision of Dr. Marielle Simon. I am conducting a research project titled "An exploratory study of the mediating influence of extant internal accountability systems on classroom level response to external accountability initiatives". The purpose of this study is to examine the impact of accountability initiatives and policies on teacher assessment practice. The case study will explore and describe the dynamics between school and provincial accountability systems, school action plans, and how teachers cope with changing expectations. It is hoped that by incorporating teacher perspective in the exploration of these events what is learned will contribute to a better understanding of the nature of accountability systems, their influence on classroom assessment practice, and professional responses to changing expectations.

In conducting this research, your school will not be negatively impacted in any way. Indeed it is anticipated that participation in the study will heighten awareness of assessment purposes and processes while enhancing appreciation for the mediating role of internal accountability systems. If you decide to participate in this study, the researcher will visit your school prior to and following a reporting period.

During the study, the researcher will interview you, and if agreed in advance, may observe a mathematics assessment activity. The math assessment observation is optional. You will participate in one semi-structured interview and two informal interviews during the study. The initial interview will take approximately 45 minutes, the remaining ones about 30 minutes. The researcher will observe staff meetings associated with the school plan and mathematics program assessment. Informal discussion will be recorded. Teachers will be asked to provide copies of documentation, such as assessment tools, evaluation criteria, and recording and reporting procedures used in assessments. These will be photocopied and originals returned. Students' work will not be collected or used in the analysis. Condensed summaries describing the accountability initiatives (ie: school plan), associated procedures, influences and teacher assessment practices will be shared with the appropriate participant for member checking of the information. You will be invited to make comments and your input will be appreciated.

If you are not participating in the classroom teacher phase of the study, your involvement

I, _____ (print name), am interested in collaborating in the study titled "An Exploratory Study of the Mediating Influence of Extant Internal Accountability Systems on Classroom Level Response to External Accountability Initiatives" and certify that I understand the nature of this research as described above.

I agree to participate in the semi-structured interview, informal interviews, team/staff assessment meetings, and document sharing. I agree to member check the condensed summary provided.

I (am/ am not) willing to participate in a math assessment activity observation.

Teacher Signature

Date

Researcher Signature

Date

I would like to receive a summary of the research findings. _____ (initials)

Appendix B

INFORMED CONSENT
(School Administrator)

Principal investigator: Adell Hay, Faculty of Education,
University of Ottawa
Address: Rm. 369 Lamoureux Hall, Faculty of Education
Telephone: (613) 599-XXXX (h)
Email: xxxxxxxx

I am currently a Ph.D. candidate at the University of Ottawa working under the supervision of Dr. Marielle Simon. I am conducting a research project titled "An exploratory study of the mediating influence of extant internal accountability systems on classroom level response to external accountability initiatives". The purpose of this study is to examine the impact of accountability initiatives and policies on teacher assessment practice. The case study will observe and describe the dynamics between school and provincial accountability systems, school internal accountability processes, school action plans, and how teachers cope with changing expectations. It is hoped that by incorporating teacher perspective in the exploration of these events what is learned will contribute to a better understanding of the nature of accountability systems, their influence on classroom assessment practice, and professional responses to change.

In conducting this research, your school will not be negatively impacted in any way. Indeed it is anticipated that participation in the study will heighten awareness of assessment purposes and processes while enhancing appreciation for the mediating role of internal accountability systems. If you decide to participate in this study, the researcher will visit your school to meet and interview participating teachers prior to a reporting period and again following a reporting period. The visits will consist of meeting with classroom teacher participants for 3 interviews and acting as an observer at meetings which involve the school plan and mathematics assessment. The visits will all take place within a two month period and will involve approximately 6 visits to your school.

Your participation in the study would involve:

1. Granting permission for the study to be conducted in your school and informing your staff about the intent of the study.
2. Providing the names of three teachers teaching grades 4, 5, 6 and/or 7 who will be willing to participate in the classroom teacher phase of the study.
3. Providing the names of school assessment leaders who may be willing to participate in a semi-structured interview.
4. Providing the researcher with a copy of school policy and procedure manuals, school action plans, and assessment-related memorandum to staff and parents.
5. Participate in a semi-structured interview of approximately 45 minutes in length and

I, _____ (print name), am interested in collaborating in the study titled "An Exploratory Study of the Mediating Influence of Extant Internal Accountability Systems on Classroom Level Response to External Accountability Initiatives" and certify that I understand the nature of this research as described above.

I agree to grant permission for the study, provide the names of a grade 4,5,6 and/ or 7 teacher to participate in the classroom teacher phase, identify school leaders, provide needed documentation, participate in a semi-structured interview, and provide a location for the interviews to be conducted if needed. I agree to member check the condensed summary provided.

Participant Signature

Date

Researcher Signature

Date:

I wish to receive a copy of the study findings _____ (initial here)

Appendix C

INFORMED CONSENT
(School Leader)

Principal investigator: Adell Hay, Faculty of Education,
University of Ottawa
Address: Rm. 369 Lamoureux Hall, Faculty of Education
Telephone: (613) 599-XXXX (h)
Email: xxxxxxxx

I am currently a Ph.D. candidate at the University of Ottawa working under the supervision of Dr. Marielle Simon. I am conducting a project titled "An exploratory study of the mediating influence of extant internal accountability systems on classroom level response to external accountability initiatives". The purpose of this study is to examine the impact of accountability initiatives and policies on teacher assessment practice. The case study will explore and describe the dynamics between school and provincial accountability systems, school internal accountability processes, school actions plans, and how teachers cope with changing expectations. It is hoped that by incorporating teacher perspective in the exploration of these events what is learned will contribute to a better understanding of the nature of accountability systems, their influence on classroom assessment practice, and professional responses to change.

In conducting this research, your school will not be negatively impacted in any way. Indeed it is anticipated that participation in the study will heighten awareness of assessment purposes and processes while enhancing appreciation for the mediating role of internal accountability systems. If you decide to participate in this study, the researcher will visit your school to interview you and act as an observer at meetings which involve the school plan and mathematics assessment plans. The visits will all take place within a two month period.

Your participation in the study would involve:

1. Participating in a semi-structured interview of approximately 45 minutes in length and member checking the condensed summary of the interview.
2. Participating in school meetings associated with the school plan and mathematics assessment.

Interview and assessment meetings will be audio-taped and transcribed. By signing this consent form participants will be agreeing to allow the researcher to use direct quotes from the interviews, observations, and field notes of informal discussions. The identity of all participants and comments on class responses to assessment procedures will be anonymous. Pseudonyms will be used for all persons, schools and district identifiers in any publication or oral presentation resulting from this research. Personal information shared by any of the participants will remain confidential. In order to assure confidentiality, all data collected will only be available to the researcher. The data will be kept secure until the study is published, at which time the raw data

I, _____ (print name), am interested in collaborating in the study titled "An Exploratory Study of the Mediating Influence of Extant Internal Accountability Systems on Classroom Level Response to External Accountability Initiatives" and certify that I understand the nature of this research as described above.

I agree to participate in the semi-structured interview and applicable staff/team assessment meetings. I agree to member check the condensed summary provided.

Participant Signature

Date

Researcher Signature

Date

I wish to receive a copy of the study findings _____ (initial here)

Appendix D

INFORMED CONSENT
(District Level Personnel)

Principal investigator: Adell Hay, Faculty of Education,
University of Ottawa
Address: Rm. 369 Lamoureux Hall, Faculty of Education
Telephone: (613) 599-XXXX (h)
Email: xxxxxxxx

I am currently a Ph.D. candidate at the University of Ottawa working under the supervision of Dr. Marielle Simon. I am conducting a project titled "An exploratory study of the mediating influence of extant internal accountability systems on classroom level response to external accountability initiatives". The purpose of this multiple case study is to examine the impact of accountability initiatives and policies on teacher assessment practice. The study will explore and describe the dynamics between school and provincial accountability systems, the nature of school internal accountability processes, school action plans, and how teachers cope with changing expectations. It is hoped that by incorporating teacher perspective in the exploration of these events what is learned will contribute to a better understanding of the nature of accountability systems, their influence on classroom assessment practice, and professional responses to change.

In conducting this research, district schools will not be negatively impacted in any way. Indeed it is anticipated that participation in the study will heighten awareness of assessment purposes and processes while enhancing appreciation for the mediating role of internal accountability systems. If your district decides to participate in this study, the researcher will visit each school prior to a reporting period and for a second visitation following a reporting period. The study will take place over a two month period with approximately 6 visits to each school site for the purposes of conducting interviews and participating in observations of staff meetings.

Your participation in the study would involve:

1. Granting permission for the study to be conducted in your district and informing potential school principals about the intent of the study.
2. Providing the names of 4-6 schools and their principals who fulfill the criteria outlined in the request for research permission documents submitted to the Ottawa-Carleton Research Advisory Committee.
3. Providing the researcher with a copy of district policies and procedures relating to testing programs, assessment/reporting procedures, support initiatives, and EQAO district action plans.
4. Participating in a semi-structured interview of approximately 45 minutes in length and member checking the condensed summary of the interview.

The interview will be audio-taped and transcribed. By signing this consent form you will

I, _____ (print name), am interested in collaborating in the study titled "An Exploratory Study of the Mediating Influence of Extant Internal Accountability Systems on Classroom Level Response to External Accountability Initiatives" and certify that I understand the nature of this research as described above.

I agree to grant permission for the study, provide the names of 4-6 schools, supply documentation, and participate in a 45 minute semi-structured interview. I agree to member check the condensed summary provided.

Participant Signature

Date

Researcher Signature

Date

I wish to receive a copy of the study findings _____ (initial here)

Appendix E

INFORMED CONSENT
(EQAQO Representative)

Principal investigator: Adell Hay, Faculty of Education,
University of Ottawa
Address: Rm. 369 Lamoureux Hall, Faculty of Education
Telephone: (613) 599-XXXX (h)
Email: XXXXXXXXXX

I am currently a Ph.D. candidate at the University of Ottawa working under the supervision of Dr. Marielle Simon. I am conducting a project titled "An exploratory study of the mediating influence of extant internal accountability systems on classroom level response to external accountability initiatives". The purpose of this multiple case study is to examine the impact of accountability initiatives and policies on teacher practice. The case study will explore and describe the dynamics between school and provincial accountability systems, the nature of school internal accountability processes, and how teachers adjust and cope with changing expectations. It is hoped that by incorporating teacher perspective in the exploration of these events what is learned will contribute to a better understanding of the nature of accountability systems, their influence on classroom assessment practice, and professional responses to change.

In agreeing to participate in this study, you will be participating in a telephone interview which should be approximately 30 minutes in length. You are the only person being interviewed from the EQAO for this study. A synopsis of this interview will be provided to you to enable member checking of its content. By continuing with this interview you are agreeing to allow this researcher to use direct quotes from this interview, field notes taken, and written synopsis. Your identity will remain anonymous and a pseudonym will be used in any publication or oral presentation. Any personal information will remain confidential. In order to assure confidentiality, all data collected will only be available to the researcher. The data will be kept secure until the study is published, at which time the raw data will be stored in a secure location until it is destroyed. You may withdraw from the study at any time and may refuse to answer any question at any time without prejudice or penalty.

It is anticipated that findings from this research will serve to inform policy makers, administrators, educators and interested others about the impact of accountability initiatives, the nature of current internal accountability mechanisms, and related policy to practice implications. Intelligent suppositions about how school organizations and teaching professionals proactively cope with the challenges of new policies and procedures may prove useful to educators and administrators.

This research has been approved by the Research Ethics Board of the University of

Appendix F

Protocol and Questions

Phase 1 and 2 (classroom teacher, school staff)

Assessment Practice

- What are the purposes for which students are assessed and evaluated?
(diagnosis, grading, identifying special needs, behaviour management, instructional evaluation, communicating achievement expectations, assessment as a teaching strategy... assessment for learning or assessment for accountability)
- What assessment strategies are employed?
(paper & pencil tests, curriculum embedded tests, performance assessments-structured or informal, oral questioning, standardized tests)
- What is the nature of the relationship between assessment and instruction?
- What feedback is given to students?
- What characterizes the teacher as assessor?
(background, time expended, personal/professional characteristics)
- Is the teacher knowledgeable about provincial assessment policy ?
- How does this teacher access information about assessment policy?
- What does this teacher perceive to be the purpose, requirements, and procedures needed for the current assessment policy environment?
- What are the teachers concerns related to assessment?

Accountability Systems

- Does the school participate in district or provincial testing programs?
- Does this school participate in school-based testing?
- Are individuals judged by standards related to student performance?
- If yes, are the standards explicit or implicit?
- Is the school as a whole, or the principal, or individual teachers subject to any consequences if students fail to meet standards?
- To what extent have the district or provincial influenced the accountability

system of the school? (impact)

- To what extent, if any, have the district or provincial made efforts to influence the accountability system of the school? (effort)
- What assessment practices and procedures are institutionalized?
- What are the specific components of the school's internal accountability

system?

(information about the organization's performance, standards for judging quality/degree of success, significant consequences for success/ failure, agent who judges achievement and distributes rewards/sanctions)

- How well has the district implemented the standards-based assessment policy?

(To what is success attributed, What has hindered the success)

- What activities are occurring at the district level as a result of provincial

policy?

(workshop, networks, instructional resources, moderation processes, etc.)

- How well has the school implemented performance-based assessment policy?
(To what is success attributed, What has hindered the success)

- What activities are occurring at the school level as a result of district policy?
(workshop, networks, instructional resources, moderation processes, etc.)

- How has the school changed as a result of standards-based assessment policy?
(In what ways has the accountability affected the way you teach)

- What challenges have occurred in developing standards-based practices?
(Identify barriers or facilitators)

- How has the use of standards-based assessment affected the classroom?
(Think: instructional practice, student assessment and grading, classrooms resources)

Sources Phase 1:

Classroom teacher interviews, field observations, instructional & assessment materials, teacher assessments, school plans, reports and publications.

Sources Phase 2:

Interviews with principal and school leaders; school plans, newsletters, and web page; district workshops, guidelines, and publications; and field notes.

Note: Questions developed with reference to the following works: Newman et al. (1997), Keene (1995), Carter (1993), Stiggins & Conklin (1992).

Phase 3 (District administration / policy makers)

Accountability Systems

- What are the specific components of the school's [district's] internal accountability system?
(information about the organization's performance, standards for judging quality/degree of success, significant consequences for success/ failure, agent who judges achievement and distributes rewards/sanctions)
- What assessment practices and procedures are institutionalized?
- Do the schools participate in district or state testing programs?
- Are individual schools/teachers judged by standards related to student performance?
- Are school/district standards explicit or implicit?
- Is the school as a whole, the principal, or individual teachers subject to any consequences if students fail to meet standards?
- To what extent, if any, have the district or state influenced the accountability system of the school?
- To what extent, if any, have the district or state made efforts to influence the accountability system of the school?
- What is the perceived assessment policy environment?
- How are assessment policies and procedures communicated with the schools?
- How much influence do schools have over content of in-service programs, assessment procedures and reporting decisions?
- What influences do professional organizations (i.e. The Teachers' College), Ministry personnel, and other stakeholders have on teachers and schools related to assessment literacy and issues of assessment practice?

Sources Phase 3:

Interviews with program directors, board policy directives, provincial directives, workshops, workshops, district strategic plans, district website.

Phase 4 (EQAO)

Accountability Systems

- What are the specific components of the province's external accountability system?

(information about the organization's performance, standards for judging quality/degree of success, significant consequences for success/ failure, agent who judges achievement and distributes rewards/sanctions)

- What assessment practices and procedures are institutionalized?
- Do all schools /districts participate in provincial testing programs?
- Are individual districts /schools /teachers judged by standards related to student performance?
- Are school/district standards explicit or implicit in district action plans?
- Is the district or the school as a whole, the principal, or individual teachers subject to any consequences if students fail to meet standards?
- To what extent, if any, have the province influenced the accountability system of the district /school?
- To what extent, if any, has the province made efforts to influence the accountability system of the school?
- What is the perceived assessment policy environment?
- How are assessment policies and procedures communicated to the district and the schools?
- How much influence do schools have over content of the tests, in-service programs, assessment procedures and reporting decisions?
- What influences do professional organizations (The Teachers' College), Ministry personnel, district staff, teachers and other stakeholders have on EQAO policy and procedures ?

Source Phase 4:

Interview with a representative of EQAO, EQAO publications and website.

APPENDIX G

Transcription/Data Retrieval Start Codes**Accountability Systems**

IntAcct	sc (scores, data on performance)
	st (standard)
	st (standard used, explicit)
	expec (collective expectations, implicit)
	ag (agent(s) to judge performance)
	cq (consequences, mechanisms)
	SIP/DIP (goals, strategies, processes)
	goals (goals and aims)
	target (target-setting exercise)
	TPA (annual learning plans)
	teacherplans (long range plans)
	repu (school reputation)

Factors (avenues of influence)

Factors	res (human, technical, fiscal)
	res (human, curricular, fiscal)
	time (pressure)
	stress (anxiety, student needs)
	lit (assessment literacy)
	pd (professional development, learning opportunities)
	cc (contextual)
	struc (organizational structure)
	pc (professional community)
	lead (leadership)
	coll (collegiality)
	own (ownership, values)
	quo (status quo, resistance)
	union (political issues)

Accountability Framework

AcctFr	mc (mandated change)
	tri (triggers for change)
	EQAO (requirements)
	reportcard (standardized requirements)

si (self initiated)
lc (locus of control -decision-making)
EQAO (policy aims, instruments, strategies)
DIP (aims, goals, instruments, strategies)
SIP (aims, goals, strategies, instruments)
al (coherence of policy aims, instruments, strategies)

Mediating Influence on Assessment Practice

MI **tools** (assessment instruments)
pur (purpose)
asstpr (methods, procedures)
eval (evaluation)
prog (programming changes)
reorg (organizational changes)
time (time allocation)
commun (communication, feedback)

Reorganization of Codes

Accountability Systems

IntAcct **goals** (aims and goals)
sc (scores, data on performance)
st (standard)
st (standard used, explicit)
expec (collective expectations, implicit)
ag (agent(s) to judge performance)
cq (consequences, mechanisms)
SIP (goals, strategies, processes)
prog (programming changes)
reorg (organizational changes)
time (time allocation)
target (target-setting exercise)
TPA (annual learning plans)
teacherplans (long range plans)
reportcard (standardized requirements)
eval (program review and evaluation)

Factors (avenues of influence)

Factors	EQAO (requirements, instruments, strategies)
	repu (school reputation)
	lit (assessment literacy)
	lead (leadership aims, strategies)
	res (human, technical, fiscal)
	time (pressure)
	stress (anxiety, student needs)
	pd (professional development, best practices)
	cc (contextual)
	struc (organizational structure)
	pc (professional community)
	coll (collegiality)
	quo (status quo, resistance)
	union (political issues)

Accountability Framework

AcctFr	mc (mandated change)
	tri (triggers for change)
	EQAO (requirements)
	reportcard (standardized requirements)
	si (self initiated change)
	lc (locus of control -decision-making)
	EQAO (policy aims, instruments, strategies)
	DIP (aims, goals, instruments, strategies)
	SIP (aims, goals, strategies, instruments)
	al (coherence of policy aims, instruments, strategies)

Mediating Influence on Assessment Practice

MI	asstpr (assessment practice)
	tools (assessment instruments)
	pur (purpose)
	proced (methods and processes)
	commun (feedback, parent and public communication)
	own (ownership, values)
	quo (status quo, resistance)

Dimensions		Guiding Questions	Elements
Conceptions underlying assumptions		<i>What beliefs and ideas about accountability are held?</i> <i>What are the key sources of authority?</i>	Accountability policy can influence schools to accomplish <u>goals</u> through: compliance, competition, democratic principles, & / or professionalism. Accountability <u>mandates</u> and <u>processes</u> can be aligned to improve school performance through: hierarchical structures, external oversight, or local management. Bureaucratic, legal or moral <u>sources</u> of authority.
Approach mode		<i>What is to be accounted for?</i> <i>What level of accountability is to be provided?</i>	Organizational effectiveness Organizational efficiency Standards of practice Standards of moral behaviour Best practices Organizational change process School ranking / rating Description, explanation & / or justification
Aims & Consequences		<i>What purpose does this policy serve?</i> <i>What are the consequences of providing an account?</i>	To change teachers' practice To stimulate accountability To enhance student achievement <i>Voluntary</i> reporting. If no obligation to report, no accountability exists <i>Obligatory</i> reporting but no consequences specified (moral / social obligation) <i>Obligatory</i> reporting with specific rewards or sanctions
Form policy		<i>What instruments and strategies are used to achieve the aims?</i>	Curriculum framework Achievement standards Mandates (e.g., testing programs) Certification (e.g., Standards of Practice) System change (e.g., structure, relationships) Inducements (rewards, sanctions, motivation)
Factors sphere of influence		<i>What avenues of influence enhance the implementation or inhibit it?</i>	Supportive actors mobilized (leadership) Professional networks (e.g., interest groups) Resources/supports (political will, funds) Requirements & / recommendations
Context & Conditions nested	Locus of Authority decision-making level	<i>At what level are the decisions made?</i> <i>To whom is the account owed?</i>	Provincial policymakers District administration School administration & teachers Public, School Boards, Parent Councils, Administrators
	Organizational Involvement depth and breadth	<i>What person(s) are involved in the response?</i> <i>Who provides the account?</i>	District administration, principals, teachers, students, parents Districts, Schools, principals, teachers

Figure H1. Dimensions of an Accountability Framework.

Note. This chart represents an interpretation of the conceptual dimensions of an accountability framework based on the literature in the field of educational administration and policy (Adams & Kirst, 1996; Claussen, 2001; Leithwood & Earl, 2000; Leithwood et al., 1999).

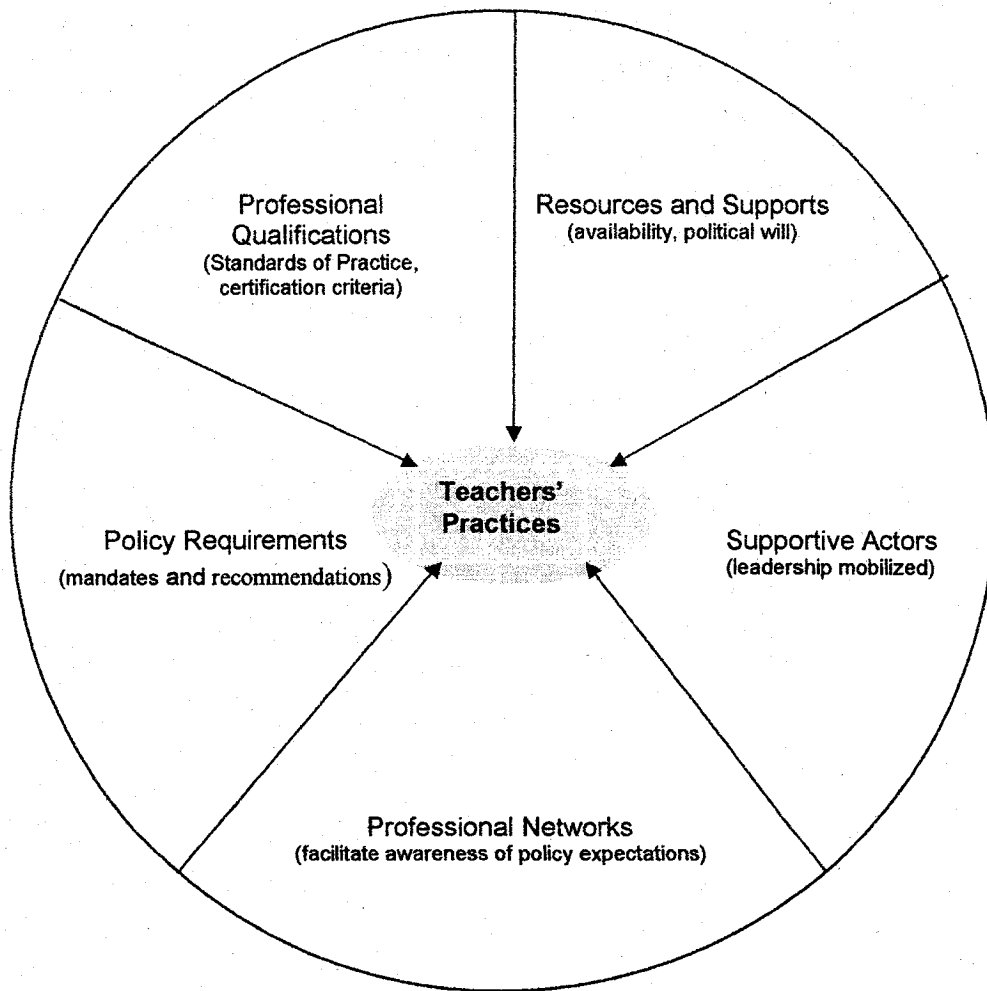


Figure H2. Avenues of Influence: Factors Influencing Internal Accountability Systems

Note. Referenced from *Between systemic reform and the mathematics and science classroom: The dynamics of innovation, implementation and professional learning*, by M.Knapp, 1997, *Review of Educational Research*, 67 (2), 227-266.

Characteristics	Classroom Assessment
<i>Purposes</i>	Diagnostic assessment- assessment carried out prior to instruction to determine student's attitudes, skills or knowledge
	Formative assessment- assessment designed to provide direction for improvement or adjust a program to meet student needs
	Summative assessment- assessment designed to provide information to be used in making a judgement about student achievement
	Student management- use of assessment to motivate or manage student behaviour
	Test-taking preparation- assessment for the purpose of exposing students to testing formats
	Accountability- assessing students for purposes of system accountability
<i>Criteria for selection</i>	Methods reflect intended outcome- the assessment is selected based on the knowledge and skills taught (content knowledge, procedural knowledge, application in problem-solving, or communication reasoning). The selection is made after the fact, or modified after the instruction.
	Adequate items/objectivity – the assessment possesses sufficient number of questions to gauge students achievement and is comprehensive in terms of the knowledge outcomes for the unit of study.
	Origin of the assessment – the assessment is useful because of its source.
	Strategy and product match purpose – the type of assessment selected suits the purpose for which the assessment was conducted (if formative it assists with scaffolding learning, process versus product).
	Economy of use – the assessment is easy to use and score.
<i>Assessment Methods</i>	Standardized tests – psychometric tests designed by measurement experts.
	Classroom recitation – question and answer, interviews, conversations and comments from observers.
	Text embedded – test questions located in textbooks or included in the instructional materials available to the teacher. May include quizzes, tests and projects.
	Performance assessment: <i>Formal</i> – structured and formal performance assessments (defined reason, pre-planned exercises, pre-specified response, scoring procedures). May include quizzes, tests and projects. <i>Informal</i> – spontaneous classroom event based on observation and judgement of students' performance or proficiency (reason not pre-defined, usually diagnostic for programming purposes, assessment was not anticipated).
	Portfolio collections – a systematic collection of students' work that represents a range of learning activities and skills students are engaged in and enables the development of student self-assessment.
<i>Feedback</i> -oral -non-verbal -written	Descriptive <u>versus</u> symbolic – written comments that are oriented to the task as compared to praise and supportive spoken statements or no feedback.
	Guidance <u>versus</u> praise – feedback is directed to the task as compared with praise directed to the student (self) away from the task.
	Germane <u>versus</u> irrelevant – feedback that is task specific and related to criteria and expectations (anticipated outcome).
	Immediate <u>versus</u> delayed – immediacy to assist with scaffolding present learning to new learning.
	Focus on process <u>versus</u> product - feedback that focuses on the learning process linked with higher motivation and outcomes for most students as compared with feedback on the final product.
	Teacher assessment <u>versus</u> self/peer assessment – acceptance of others' assessment of achievement versus self-monitoring intrinsic to reflecting on one's work where the individual learns to better approximate their perception of their own work.

Figure H3. Classroom Assessment Characteristics.

Note. These characteristics are presented here to differentiate assessment practices based on the paradigm shift from a culture of testing to a culture of assessment (Biggs, 1995; Gipps, 1994).

Characteristics	Individualistic assessment	Collegial assessment
Goals	The goals reflect reluctant or superficial compliance of assessment policy.	The goals reflect compliance and acceptance of assessment policy.
Instruments achievement level portfolio exemplars	Achievement level assigned by personal interpretation of criteria.	Achievement chart criteria discussed collaboratively
	Portfolios not used.	Active use of portfolios.
	Teacher may use commercially produced tests aloof from curricular instructional activities.	Teachers own and use school and system exemplars. Embedded assessment, no commercial products.
Methods diversity of assessment strategies and purposes language of assessment moderation activity	Minimal collection of evidence to demonstrate student learning. Assessment is product oriented and measurement. Grades and use of praise common.	Assessment collections are planned, integral to instruction, employ a toolkit (several strategies), focuses on process goals as well as end product. Diagnostic, formative and summative evidence collected. Feedback positive and ongoing.
	Lacks understanding of assessment terminology, confusion of meaning.	Uses correct assessment terminology to communicate.
	No collaboration to develop consistency in assessment practice.	Discussion of assessment criteria and collaborative development or use of assessment tasks and tools (e.g., rubrics).
Personnel school-wide involvement involvement of students and parents	Individuals feel alienation from professional community. -contrived collegiality or balkanization in school assessment practice. -lack of discussion on differing perspectives	School-wide involvement in the development of a consistent approach to assessment. -regular review of school, district and state scores. -identification of strengths and weakness -shared decision-making -professional learning (best practices promoted) -district planning documents reviewed -consultation with parents and community
	Assessment is individualistic.	Assessment involves teachers, colleagues, students and parents.
Value system ownership collaborative student-oriented	Teacher works independently. -assesses for grading purposes -measurement focuses on the outcome (end product)	Teachers value collaborative process. -understand role of assessment in improvement of instructional practice -articulate benefits of student assessment -articulate benefits of learning skill (process) focus
	Teacher does not seek support.	Principal and teachers seek support and resources from district and local community (parents, parent's councils, community groups).
	Classroom assessment focused on paper and pencil tests. -summative assessment -content and procedural knowledge focus -results for grading purposes -feedback (symbolic, praise is self-oriented)	Classroom assessment focuses on performance tasks and assessments. -formative assessment focuses on process learning -feedback (comments, fair, germane, immediate) -variety of strategies to take into account student learning styles -results used to develop individual learning plans / to inform instructional decisions

Figure H4. Characteristics of School (Teachers') Assessment Practices.

Note. From "Level descriptions and teacher assessment in England: Towards a community of assessment practice, by K. Hall & A. Harding, *Educational Research*, 44 (1), p. 6. Copyright 2002 by Educational Research, <http://tandf.co.uk/journals>. Adapted with permission.

Dimensions	Modern Reform	Restructuring Systems	
		Bureaucratic control (New)	Professionalism (emerging)
Conception -assumptions -beliefs	Mandate school changes to accomplish outcomes -bureaucratic and legal authority	Hierarchical organizational structures aligned to accomplish school goals via mandates -bureaucratic, legal and moral authority	Democratic processes of education can be aligned to accomplish school goals through professional empowerment -moral sources of authority
Approach <u>What is accounted for</u> <u>What intensity of accountability</u>	<u>Systemic Reform Policies</u> - approved models of instruction -measured outcomes -changes focussed at school level (description or explanation)	<u>Adam & Kirst's (1996) Classification scheme</u> 1. School-based management 2. School choice 3. Teacher empowerment 4. Teaching for understanding (description, explanation, or justification)	<u>Leithwood et al.'s (1999) Classification scheme</u> 1. Management 2. Market competition 3. Decentralized decision-making 4. Professionalism a) professional control b) professional practice (description, explanation and justification)
Aims & Consequences	-change school programming -use outcome measures to evaluate school performance -rewards based on mastery of basic skills (extrinsic motivation)	- use incentives (rewards, sanctions) to motivate school performance and organizational change -rewards based on students' performance (extrinsic motivation)	-establish a forum for accountability and teacher learning - professional ethics and professional esteem (intrinsic motivation)
Instruments & Processes	- new curriculum and program materials - standardized multiple choice testing -policy mandates -teacher certification	- curriculum expanded to higher skills -criterion-referenced performance testing -policy mandates -requirements -teacher re-certification	-curriculum (with standards) -criterion-referenced performance/multiple choice testing method -policy menus -recommendations -teacher re-certification
Factors	-resources -teacher training	-resources (human, technical, fiscal, political) -professional development	-resources (human, technical, fiscal, political) - professional learning
Decision-making Authority	-trust in authority -centralized decision-making - external accountability - mandated consequences	-trust in people -decentralized decision-making - external accountability - mandated consequences	- trust in process - responsible decision-making - internal accountability - mandated processes
Organizational Involvement	-hierarchical structure -teacher as service provider -principal as manager -system accountability	-collective vision -organizational culture -teacher professional role differentiated -broadened school accountability	-collaborative voice -collaborative culture -teacher role further differentiated -principal as instructional leader -community involvement in accountability

Figure H5. Continuum of Approaches to Accountability.

Note. This chart represents an interpretation of the relationship between different philosophical approaches reported in the literature (Adams & Kirst, 1996; Earl & Torrance, 2000; Fuhrman, 1999; Hargreaves, 1996; Leithwood & Earl, 2000; Leithwood et al., 1999; Mintrop et al., 2001).

Dimensions	Southview School
Aims and Consequences	<u>Goals of Principal's Strategic Plan</u> - conduct School Review audit to identify strengths and weaknesses using principles of effective schools and assess what aspects of teaching, learning and operation need support - effective schools strategies incorporated - school improvement planning (Literacy & Numeracy initiatives) - staff training, development and support - expectations for learning/accountability - promotion of student-oriented strategies The plan is discussed with the superintendent. Feedback received. Principal collaborates with other principals on plan implementations issues. <u>Goals of Southview SIP</u> - improve students' levels of achievement in writing - enhance teacher instruction and student achievement in mathematics - maintain a safe, caring school & include community involvement Parent Council, community leaders and teachers are consulted during plan development. The plan is shared with the Parent Council and Superintendent. Feedback is received on program effectiveness from EQAO (scores), district administration (statistical analysis), parent council (expectations) and community leaders (inter-school continuity). School reputation may impact on market share. <u>Consequences:</u> - report, explain, justify instruction program (school reputation, professional esteem, market forces [e.g., school choice])
Information on school performance	School collects information on students abilities and achievement from: - EQAO tests at Grade 3 and 6 - Grade 7 mathematics diagnostic test - Early Development Instrument (EDI) & Canadian Cognitive Abilities test (CCAT) - Teacher assessments and grades - IEPs - EQUIP & school community data - Parent Council and community leader feedback
Standard of Achievement	School standard - provincial standard is school target (promoted by principal and parent community) - teachers identify with provincial standard for strong students, but express concern for weak students (special needs, ESL)
Persons who judge school performance	<u>Teaching team level:</u> - examine EQAO results and trends in the data - reflect on school performance taking all current information on student population into consideration - Grade 3 and 6 teachers examine class data and reflect on students' performances against own assessment of student achievement (calibrate standards) <u>School level:</u> - administrative team led by principal examines school data (school audit information, diagnostic and achievement test results, report card grades) and district statistical analyses (EQAO school & district aggregated results) - administrative team consults with quality assurance staff and superintendent - principal consults with resource teacher and classroom teachers <u>Community level:</u> - principal consults with Parent Council and community leaders (High School principal and parent representatives) to clarify expectations and priorities

Figure H6. Southview School Internal Accountability Components.

Note. This chart displays data collected and reduced from district and school interviews, documents and field notes.

Mechanisms to guide & monitor teachers' practices	<u>Provincial instruments & strategies:</u> - EQAO performance-based testing program - public reporting of district and school results - district and school improvement plans (DIP & SIP) - standardized report card <u>Provincial Feedback:</u> - district and school aggregated scores, individual students scores - recommendations for district and school actions <u>Other instruments:</u> target setting, provincial exemplars, curriculum, teacher standards of practice (College of Teachers), special funding projects (incentives)
	<u>District instruments:</u> - School review guidelines - School improvement plan guidelines - Assessment and grading guidelines <u>District strategies:</u> - SIP workshop (Quality Assurance Department input) - Staff training and development - Expectations for learning and accountability (district literacy, numeracy & safe schools initiatives) - Community partnerships <u>District feedback:</u> - Superintendent monitors strategic plan and SIP - professionalism, intrinsic rewards
	<u>School instruments:</u> - Principal's strategic plan including the SIP, School Report, - Report card writing process - Annual professional growth plan - Long range plan template <u>School strategies:</u> - administrative team evaluation of school performance - collegial dialogue promoted - professional development (best practices introduced) - curriculum planning structured at working staff meetings (team planning) - replacement of instructional resources - resource inventories and centralized storage - primary, junior and intermediate team reorganization (teacher reassignment by principal) - timetabling changes (to promote team teaching, increase instructional time) - long range planning (principal supervision & support) - professional growth linked to SIP plans (principal supervision) <u>School feedback:</u> - program evaluation and feedback from professional reflection, collegial dialogue, principal consultations, and community input (school reputation) - teacher performance appraisals (intrinsic rewards, collegial esteem) <u>Other instruments:</u> Provincial curriculum documents and supplementary resources, provincial achievement chart & rubrics (standard), standardized report card software program, EQAO testing anchors and sample tests.

Figure H6 (con.). Southview School Internal Accountability Components.

Note. This chart displays data collected and reduced from provincial, district and school interviews, documents and field notes.

Factors	<i>Southview School</i>
<u>External requirements</u>	EQAO performance-based testing program: - provided a measure of school performance against a provincial standard of achievement - required schools & districts to publish school, district and provincial scores - required school improvement planning that explained trends in school data, justified goals and strategies for change, and involved school, district & community consultations
<u>Mobilization of Supportive Actors</u>	The principal's leadership enhances the internal accountability system by actions listed: - conducted a school review in consultation with teachers, parents and community leaders - developed a strategic plan that addressed 6 core factors associated with effective schools - used a professional approach to school restructuring and shared her personal vision and mission with staff - focussed on student achievement and a supportive school environment - analyzed a variety of data on school performance (school, district, provincial) and set goals in consultation with district administration and school leaders - involved teachers in the decision-making stage related to instructional strategies - owned school organizational changes and reassigned teaching staff based on student programming needs - facilitated professional development and collegial sharing of best practices - assumed a mentor role with teachers and held teachers accountable for meeting personal professional growth goals - teacher initiative to arrange own professional development - limited time for professional reflection, collegial planning of performance tasks/assessments, professional development and collegial dialogue
<u>Political Will</u>	Political Issues inhibited internal accountability because: - school board in non-compliance with government due to budget deficit - teacher unions sanctioning EQAO marking sessions and teacher re-certification program (PLP courses)

Figure H7. Factors Impacting Southview School's Internal Accountability System.

Note. This chart displays data collected and reduced from provincial, district and school interviews, documents and field notes.

EQAO*	District	Southview School
Aims		
a) assure greater accountability and contribute to the enhancement of quality education through assessments and reviews based on objective, reliable and relevant information (EQAO tests scores, EQUIP).	a) direct operational planning to support student learning; inform Board policy and budget decisions; and focus improvement efforts of schools and departments	a) conduct a school review to identify strengths and weaknesses in the teaching, learning and operations of the school and develop a SIP that focuses on district literacy, numeracy and caring schools initiatives appropriate to identified school priorities
b) make recommendations for system improvement and monitor system change through feedback; provide system transparency through public reporting.	b) explain objectives and decisions to parents and the public; provides explanation to the EQAO through submission of DIP	b) consult, explain and justify program effectiveness, improvement strategies and expectations with district administration, school staff, parents and the community
Information on School Performance		
Student EQAO test scores, Data on the learning environment inside and outside school (EQUIP.)	Standardized testing (CCAT), EDI, Grade 7 diagnostic math test, school community demographics, program characteristics.	Special needs testing (SERT), classroom assessments and grades, school review findings community input, parent council expectations & priorities.
Standard of Achievement		
Provincial achievement chart (standard)	Provincial achievement chart (standards) District guidelines on assessment & grading practices developed.	Use of provincial achievement chart (standards), district guidelines, school/community expectations, IEPs where appropriate.
Person(s) Judging School Performance		
EQAO publishes performance results, makes recommendations to stakeholders, provides feedback on DIPs.	Superintendent monitors Principal's school review audit & strategic plan, including SIP. Quality Assurance personnel provides input into SIPs (statistical analysis, EQUIP, socio-economic & demographic data). District staff publish and distribute the School Reports (SIP& DIP) to public. Board reviews System Review audit	Principal conducts school review. Principal mentors teacher personal growth plans and monitors instructional program. Principal consults with Parent Council, superintendent & community leaders. Principal & administrative team consult with district experts to evaluate school performance and set goals for SIP. School staff reviews school performance, owns SIP goals and plan next steps. Principal distributes School Profile and student EQAO results to parents.

Figure H8. Relationship between External and Internal Accountability Systems.

Note. This chart displays data collected and reduced from district and school interviews, documents and field notes and from the EQAO Provincial Report on Achievement (1997).

Mechanisms		
<u>Provincial instrument:</u> - EQAO performance-based testing program - EQUIP <u>Provincial strategies:</u> - district and school EQAO results reported to public - district and school improvement plans required (DIP & SIP) <u>Provincial Feedback:</u> - reports district and school aggregated scores, individual students scores - makes recommendations for district and school actions <u>Other instruments:</u> target setting, provincial exemplars, anchor books, standardized report card and reporting requirements, curriculum and provincial achievement standards, teacher standards of practice (College of Teachers), special funding projects (incentives)	<u>District instruments:</u> - District strategic plan (system renewal) - School review audit guidelines - District testing program & school demographic data - School improvement plan guidelines - Guidelines on use of statistical data - Assessment and grading guidelines <u>District strategies:</u> - Superintendent supervision of school review process - SIP workshop (Quality Assurance expertise & input) - Staff training and development - District literacy, numeracy & safe schools initiatives - Community partnerships <u>District feedback:</u> - Superintendent monitors strategic plan (actions to address 6 factors of effective schools) - Superintendent consults on SIP <u>Other instruments:</u> superintendent visitations, administrative meetings and workshops	<u>School instruments:</u> - Principal's strategic plan - SIP, School Report - Long range plan template <u>School strategies:</u> - Principal conducts school review and consults with superintendent, colleagues, parents and community about strategic plan - Principal revisits school mission statement and consolidates the collective vision - Principal reviews school performance and consults with administrative team to set instructional priorities (SIP) - team planning promoted (professional dialogue) - professional development (best practices introduced) - working staff meetings (team co-operation and planning) - replacement of instructional resources - resource inventories & centralization - primary, junior and intermediate team reorganization (teacher reassignment by principal) - timetabling changes (to enable regrouping & team teaching, increase instructional time) - long range planning (principal monitored, supports) - professional growth plans (criteria set and checked, classroom visitations) <u>School feedback:</u> - program review and feedback from EQAO results, professional dialogue, principal consultations, and community input - teacher performance appraisal (intrinsic rewards, collegial esteem/sanction)

Figure H8 (con.). Relationship between External and Internal Accountability Systems.

Note. This chart displays data collected and reduced from district and school interviews, documents and field notes and from the EQAO Provincial Report on Achievement (1997).

General Recommendations Specific activities for Schools.	Southview School Response
Communicating and interpreting findings 2. Schools consult with teachers, parents and the community to establish actions plans and review them on a regular basis. 5. Classroom assessment strategies to find out where students are in their learning and what they need to do next. 35. Identify school and community factors related to student achievement. 36. Communicate about the purposes of assessment, the expectations for students, the quality of student work and the ways that the school and the home provide support for learning. 37. Discuss achievement and improvement strategies.	2. – implementation is established, the process and plan are effective, improvements are continuing. [SIP process, School Review audit, School Report, consultation with parent Council and Community stakeholders] 5. – practice is common and variety of strategies being implemented, difficulties have been identified and improvement initiatives are being carried out. [performance based tasks, formative assessment in textbook series, target setting, preparing for the test] 35. - school audit completed and improvement strategies in place for instructional program and school operations, implementation is being monitored by principal and superintendent, results are documented and communicated to stakeholders. [School Review audit, SIP, EQUIP and socio-economic information gathered from consultations, Community input] 36. – principal and teachers using assessment toolkits, difficulties identified and improvement initiatives being carried out to involve parents in extra practice/assessment discussions, student taught expectations and evaluative criteria. [student assessment (curriculum expectations and criteria taught), homework help, rubrics and checklists, Parent Council dialogue] 37. – high levels of implementation and effectiveness are evident, implementation is monitored, results documented. [SIP, report card, parent-teacher interviews, IEP, consultations with community stakeholders, School Report]

Figure H 9. Southview School Response to Accountability Aims.

Note. This chart displays data collected and reduced from school interviews, documents and field notes and from the EQAO Provincial Report on Achievement, 1997.

Characteristics	MODEL	Southview School
Purposes	Diagnostic assessment	Student knowledge assessed by teacher to assist with instructional planning prior to unit of study. Common practice, may be oral discussion or a quiz
	Formative assessment	Ongoing assessment of student learning skills throughout unit, including some journaling. Feedback by teacher to improve performance a common practice. (oral)
	Summative assessment	Overall performance (achievement & learning skills) evaluated using curriculum expectations and standards (criterion-referenced) was consistent practice. Final grade based on best, most consistent performances throughout the unit of study.
	Student management	Assessment used as feedback to motivate students was common practice.
	Test-taking preparation	Performance tasks and assessment emphasized before EQAO by Grade 6 teacher to provide feedback on learning skills.
	Accountability	School performance assessed by EQAO assessments and teacher grades to assist with program assessment (principal & administrative team & public). Student performance was self/peer assessed to develop self-assessment skills and enhance learning (feedback).
Criteria for selection	Methods match the outcome	Some assessments planned after instruction is completed and designed to assess what was taught. Math journaling to assess communication and reasoning. Performance tasks selected to assess procedural knowledge.
	Adequate items/objectivity	Additional questions added to textbook assessments to fulfill curriculum outcomes. Textbooks assessments considered objective. Subjective component of assessment recognized.
	Origin of the assessment	Assessments from approved textbooks and teachers resources in common practice. Assessment from Ontario Curriculum planner, exemplars and EQAO anchor books an occasional practice. Some teacher constructed tests designed to match expectations.
	Results fit the purpose	Performance tasks used to motivate students and scaffold learning. Math journaling and demonstrations used for summative assessment purposes.
	Economy of use	Textbook materials possess a variety of prepared assessments and tools (rubrics, checklists). Use is common practice but may be modified.
Method of assessment	Standardized tests	District wide testing, Special Education testing common practice.
	Classroom recitation	Oral questioning by teacher common practice for assessment and for learning (scaffolding).
	Text embedded	Mathematics program possessed extensive examples. (common practice)
	Performance assessment -formal -informal	Performance assessments planned in advance is common practice. Some performance tasks developed by teachers, others found in documents. Informal teacher performance assessment is common practice.
	Portfolio collections	Common as part of the writing process, not in mathematics.
Feedback -oral -non-verbal -written	descriptive vs symbols	Some written comments but usually linked with letter grade or performance level. Rubrics and checklists in common use.
	guidance vs praise	Praise in common use, but feedback beginning to focus on what students can do and next steps in the some classrooms.
	fair vs unfair	Feedback related to criteria, transparent processes improving, no surprises. Oral feedback on process of learning often informal.
	germane vs irrelevant	Focus on performance is emerging, but unit end tests focus on product. Increased focus on assessing to criteria, but subjective aspects remain when recognizing effort/ commitment to task.
	immediate vs delayed	Written feedback and grades sent home to parents, enlisting support in some junior classes.
	process vs product	
	parental communication	

Figure H 10. Classroom Assessment

Note. This chart displays data collected and reduced from school interviews, documents and field notes.

Characteristics	Southview School
Goals - acceptance vs resistance	School goals reflect compliance with and acceptance of standards assessment policy, initial resistance by individuals changing.
Instruments -collective criteria -portfolio use -exemplars in use	Assessment criteria based on the expectations and outcomes of the curriculum and assessed against the levels of performance described in the provincial achievement charts. Use of district Assessment and Grading handbook. Portfolio use in writing only, student life skills portfolio a goal of principal's strategic plan. Provincial exemplars and EQAO anchors used in classroom assessment.
Methods (Processes) -planned collection -embedded -variety of strategies (toolkit) -process goals -diagnostic, formative & summative assessment -assessment literate -collaborative development of tasks and tools	Planning assessments as part of unit development common but not consistent practice. Some assessment decisions remain post instruction. Use of embedded assessments a function of the new mathematics textbook program. Common but not consistent practice. Assessment toolkit evident by variety of assessment strategies used. Use of performance assessments common practice at junior and intermediate levels. Diagnostic, formative and summative assessments are common in mathematics. Feedback provided to students. Assessment literacy developing but range of skills evident. Sharing of performance-based tasks and assessment tools done informally, best practices being promoted, team planning and moderation of evaluative criteria being structured but not spontaneous at junior and intermediate level.
Personnel -school-wide involvement -student and peer -parent and community	School wide involvement in improvement planning (SIP) at different levels of responsibility. School review, goals and program changes primarily administrative decisions. Support for using new assessment strategies. Student self-assessment a common practice but not consistent among all teachers. Parent and community participate in SIP process. Parental involvement in classroom assessment processes encouraged.
Values -teaming valued -appreciates assessment as instruction -uses student self-assessment -learning skill (process) focus -assessment addresses learning styles -results used for individual student planning -consultation with community	Team approach to assessment not yet common practice at the junior and intermediate levels. Some interest expressed. Assessment articulated as part of instruction a common practice. Benefits of student self-assessment articulated. Performance tasks common practice, increased focus on assessing learning processes in addition to assessment final product. Assessment involves a variety of methods but paper and pencil unit end tests remain common practice suggesting an added on perspective. Diagnostic and formative assessment common but not consistent practice. Use of diagnostic and formative assessment to plan for individual student instruction is informal practice at junior and intermediate levels, except for special education students (IEP'd). School administrators discuss needs with district staff, parent council. Principal consults with members of the community and feeder schools. Parental support for homework and extra practice common at primary and junior level.

Figure H 11. Dynamics of Southview School Assessment Practices.

Note. This chart displays data collected and reduced from district and school interviews, documents and field notes.

Dimensions	Northside School
Aims and Consequences	<u>Goals of Principal's(SBDM Committee) Strategic Plan</u> - conduct school review to identify strengths and weaknesses in a number of areas and assess what aspects of teaching, learning and operation are in need of support - effective schools strategies incorporated - school improvement planning (School Climate, Literacy & Numeracy initiatives) - staff training, development and support - consult on school vision & develop mission statement The plan reviewed with the superintendent. Feedback on effectiveness. Administration collaborates with other Signal School principals. <u>Goals of Northside SIP</u> - improve % of students performing at or above level 3 in Reading (Grade3 target is 30%) - enhance teacher instruction and student achievement in mathematics - implement School Climate change plan (physical environment, code for learning, inclusionary education, community relations) - district experts, other signal schools, teachers & parents council are consulted during plan development. - feedback on program effectiveness from EQAO (scores), district administration (statistical analysis), signal school network(strategies), parent council (expectations) and community support groups. Parental involvement limited due to language and cultural issues. <u>Consequences:</u> - report, explain, justify instructional program (school reputation, professional esteem)
Information on School Performance	School collects information on students abilities and achievement from: - EQAO tests at Grade 3 and 6 - Grade 7 mathematics diagnostic test - Early Development Instrument (EDI) - Canadian Cognitive Abilities test (CCAT), Canadian Achievement Test (CAT) - Teacher assessments and grades - IEPs - EQUIP & school community data - Parent Council and community leader feedback
Standard of Achievement	School standard - provincial standard is school target (promoted by principal and district administration) - teachers identify with provincial standard for strong students, but express concern for affective issues with ESL & L.D. students (also truant & transient students)
Persons who judge School Performance	<u>Teaching team level:</u> - examine EQAO results and trends in the data - reflect on school performance taking all current information on student population into consideration - Grade 3 and 6 teachers examine class data and reflect on students' performances against own assessment of student achievement (calibrate standards)
	<u>School level:</u> - SBDM Committee mentored by principal examines school data (school audit information, diagnostic and achievement test results, report card grades, demographic information) and district statistical analyses (EQAO school & district aggregated results) - SBDM Committee consults with teachers, quality assurance staff & superintendent - principal consults with superintendent, resource teacher
	<u>Community level:</u> - principal consults with Parent Council, Signal Schools and community support groups to identifies priorities and seek supports

Figure H 12. Northside School Internal Accountability Components.

Note. This chart displays data collected and reduced from provincial, district and school interviews, documents and field notes.

Mechanisms to guide & monitor teachers' practices	<u>Provincial instruments & strategies:</u> - EQAO performance-based testing program - public reporting of district and school results - district and school improvement plans (DIP & SIP) - standardized report card <u>Provincial Feedback:</u> - district and school aggregated scores, individual students scores - recommendations for district and school actions <u>Other instruments:</u> target setting, provincial exemplars, curriculum, teacher standards of practice (College of Teachers), special funding projects (incentives)
	<u>District instruments:</u> - School review guidelines - School improvement plan guidelines - Assessment and grading guidelines <u>District strategies:</u> - SIP workshop (Quality Assurance Department input) - Staff training and development - Expectations for learning and accountability (district literacy, numeracy & safe schools initiatives) - Community partnerships <u>District feedback:</u> - Superintendent monitors strategic plan and SIP - professionalism, intrinsic rewards
	<u>School instruments:</u> - Principal's strategic plan including the SIP, School Report, - Report Card writing process - Curriculum documents, exemplars, & Achievement chart (standard) <u>School strategies:</u> - SBDM Committee evaluation of school instructional program - Signal school networking promoted (professional dialogue) - professional development (best practices introduced) - team teaching, mentors, divisional & school staff meetings (team co-operation and planning) - replacement of instructional resources - resource inventories and centralization - Resource teacher and special education support teachers to assist with modified programs - long range planning (assessment variety) - professional growth plans (upgrading encouraged) <u>School feedback:</u> - SIP and program feedback from professional reflection, SBDM Committee review, principal consultations, and community input - Collegial recognition and support <u>Other instruments:</u> Provincial curriculum documents and supplementary resources, provincial achievement chart & rubrics (standard), standardized report card software program, EQAO testing anchors and sample tests.

Figure H 12. Northside School Internal Accountability Components.

Note. This chart displays data collected and reduced from provincial, district and school interviews, documents and field notes.

<i>Factors</i>	<i>Northside School</i>
<u>External Requirements</u>	EQAO performance-based testing program: - provided a measure of school performance against a provincial standard of achievement - published school, district and provincial aggregated scores - required school improvement planning that explained trends in school data, justified goals and strategies for change, and involved school, district & community consultations
<u>Mobilization of Supportive Actors</u> <u>Professional Networks</u>	The principal's leadership enhances the internal accountability system by actions listed: - conducted a school review in consultation with teachers, parents and students - developed a strategic plan that addressed 6 core factors associated with effective schools - used a professional approach to school restructuring and consulted with school community on common vision and developed a school mission statement - focussed on school climate and promoted curriculum renewal - formed a SBDM Committee to take leadership role in analysis of data on school performance (school, district, provincial) , in goal setting goals(consultation with district administration and school leader) - involved teachers in the decision-making stage related to instructional strategies - owned school organizational changes and promoted mentors and team teaching - facilitated team leadership for p.d. and collegial sharing of best practices - assumed a mentor role with SBDM Committee, team leaders and encouraged teacher professional upgrading

Figure H 13. Factors Impacting Internal Accountability at Northside School.

Note. This chart displays data collected and reduced from provincial, district and school interviews, documents and field notes.

EQAO*	District	Northside School
Aims		
a) assure greater accountability and contribute to the enhancement of quality education through assessments and reviews based on objective, reliable and relevant information (EQAO tests scores, EQUIP).	a) direct operational planning to support student learning; inform Board policy and budget decisions; and focus improvement efforts of schools and departments	a) implement 3 year strategic plan to address school climate by development of a mission statement, code of learning, & an inclusive school culture; and implement curriculum renewal by development of 2 year plan to focus on data collection, new resources, in-service, parent workshops, SIP reviews that focuses on literacy & numeracy
b) make recommendations for system improvement and monitor system change through feedback on DIP; provide system transparency through public reporting.	b) explain objectives and decisions to parents and the public; provide explanation to EQAO through submission of DIP	b) consult, explain and justify program effectiveness, improvement strategies and expectations with district administration, school staff, & parents.
Information on School Performance		
Student EQAO test scores, Data on the learning environment inside and outside school (EQUIP.)	Standardized testing (CCAT), EDI, Grade 7 diagnostic math test, school community demographics, program characteristics.	Standardized testing (SERT), classroom assessments and grades, annual CAT, school review data, parent council expectations & priorities, EQUIP & demographic data.
Standard of Achievement		
Provincial achievement chart (standard)	Provincial achievement chart (standards) District guidelines on assessment & grading practices developed.	Use of provincial achievement chart (standards), district guidelines, school/community expectations, IEPs where appropriate (approx. 50%).
Person(s) Judging School Performance		
EQAO publishes performance results, makes recommendations to stakeholders, provides feedback on DIPs.	Superintendent monitors Principal's school audit & strategic plan, including SIP. Quality Assurance personnel provides input into SIPs (statistical analysis, EQUIP, socio-economic & demographic data). District staff publish and distribute District /School profiles which include DIP & SIP to Board and public.	SBDM Committee conducts review of SIP and monitors 3 year strategic plan. Principal mentors SBDM Committee, requires personal growth plans and long range plans. Principal consults with Parent Council, district superintendent, Signal School Network & community leaders. SBDM Committee consults with district experts, evaluates school performance and set goals for SIP. School staff reviews school performance, owns SIP goals and plan next steps. Principal distributes School Profile and student EQAO results to parents.

Figure H 14. Relationship between External and Internal Systems.

Note. This chart displays data collected and reduced from district and school interviews, documents and field notes and from the EQAO Provincial Report on Achievement (1997).

Mechanisms		
<u>Provincial instrument:</u> - EQAO performance-based testing program <u>Provincial strategies:</u> - district and school EQAO results reported to public - district and school improvement plans required (DIP & SIP) <u>Provincial Feedback:</u> - district and school aggregated scores, individual students scores - recommendations for district and school actions <u>Other instruments:</u> target setting, provincial exemplars, standardized report card, curriculum and provincial achievement standards, teacher standards of practice (College of Teachers), special grants (incentives)	<u>District instruments:</u> - District strategic plan (system renewal) - School review audit guidelines - District testing program & school demographic data - School improvement plan guidelines - Guidelines on use of statistical data - Assessment and grading guidelines <u>District strategies:</u> - Superintendent supervision of school review process - SIP workshop (Quality Assurance expertise & input) - Staff training and development - District literacy, numeracy & safe schools initiatives(training workshops) - Community partnerships <u>District feedback:</u> - Superintendent recommendations for school actions based on school review (implementation of 6 factors rated) - Principal performance appraisal <u>Other instruments:</u> superintendent visitations, administrative meetings and workshops	<u>School instruments:</u> - SBDM Committee strategic plan - SIP, School report, report card - Long range plans <u>School strategies:</u> - Principal consults with staff, parents and student to develop school vision & mission statement - SBDM Committee consults with experts to review school performance and sets instructional priorities collaboratively with staff - Team teaching, mentors & divisional meetings - collegial networking promoted (signal schools) - best practices promoted - replacement of instructional resources (texts & assessment tools) - resource inventories & centralization - primary, junior and intermediate team reorganization (teacher reassignment by principal) - long range planning (principal monitored, assessment strategies) - professional growth plans (upgrading encouraged) <u>School feedback:</u> - program review(evaluation) and feedback from professional reflection, collegial dialogue, district level consultations, signal school networking and community input - teacher performance appraisals (intrinsic rewards, collegial esteem) <u>Other instruments:</u> Provincial curriculum documents and supplementary resources, provincial achievement chart (standard), standardized report card, EQAO testing anchors and sample tests.

Figure H 14 (con.). Relationship between External and Internal Systems.

Note. This chart displays data collected and reduced from district and school interviews, documents and field notes and from the EQAO Provincial Report on Achievement (1997).

General Recommendations Specific activities for Schools	Northside School Response
<i>Communicating and interpreting findings</i> 2. <i>Schools consult with teachers, parents and the community to establish actions plans and review them on a regular basis.</i> 5. <i>Classroom assessment strategies to find out where students are in their learning and what they need to do next.</i> 35. <i>Identify school and community factors related to student achievement.</i> 36. <i>Communicate about the purposes of assessment, the expectations for students, the quality of student work and the ways that the school and the home provide support for learning.</i> 37. <i>Discuss achievement and improvement strategies.</i>	2. – implementation complete, the process and plan are effective, improvements are continuing. [SIP , School Review process, School Report] 5. - a process and plan for implementing regular diagnostic testing for all students in place, difficulties identified and improvement initiatives are being carried out. [CAT testing, extra support from resource teacher(s), performance based tasks, formative assessments new instructional materials, target setting, preparing for the test] 35. – School Review audit completed and EQAO and other information screened, improvement strategies in place for inclusive school climate, renewal of instructional program, Signal School identification. [School Review, SIP, School Report, Signal School Network, community partnerships] 36. – principal and teachers using assessment toolkits, difficulties identified and improvement initiatives being carried out. Student taught expectations and criteria. Communication with parents seen as important. [rubrics, checklists, formative assessment, parent-teacher interviews, language liaison program , Parent Council dialogue] 37. – implementation of SBDM Committee plan monitored by principal and superintendent, results are documented and communicated to stakeholders, high levels of implementation and effectiveness are evident, implementation is monitored, results documented. [IEP, report cards, parent-teacher interviews, collegial assessment practice, professional learning, consultations with Parent Council and supportive health agencies]

Figure H 15. Northside School Response to Accountability Aims.

Note. This chart displays data collected and reduced from school interviews, documents and field notes and from the EQAO Provincial Report on Achievement, 1997.

Characteristics	MODEL	Northside School
Instructional purposes	Diagnostic assessment	Student achievement assessed by teachers to assist with instructional planning (prior to school year, prior to each mathematics unit).
	Formative assessment	Ongoing assessment through the mathematics unit using a variety of strategies to provide feedback on processes, progress toward the standard, to modify instruction and to plan remediation.
	Summative assessment	Overall performance evaluated against curriculum expectations or IEP criteria where appropriate. Final grade based on best, most consistent performance, including cumulative tests or performance tasks.
	Student management	Assessment used as feedback to motivate students, also used to manage student attention to task (results linked to commitment).
	Accountability	Students assessed for reporting purposes, for diagnosis of needs and assessment of program effectiveness (CAT), in preparation for EQAO Grade 6 testing.
Criteria for selection	Origin of the assessment	Assessments came primarily from the new mathematics textbook program that possessed a variety of strategies and instruments including quizzes, performance assessments, culminating activities, journal tasks and problem solving exercises. Tools provided included checklists, rubrics and anecdotal charts. Supplementary assessments were teacher developed (modified) or commercially prepared.
	Methods match outcome	Assessments selected based on expectations to be taught, planned in advance by some teachers. Modified after the fact by others based on curriculum material completed.
	Interference controlled	To reduce bias or unfairness, several small assessments may be used at intermediate stages in the learning process to ensure scaffolding of new learning. Manipulatives used as needed.
	Results fit the purpose	Math journaling used for formative assessment. Problem-solving and performance tasks for summative assessment. Teacher interview and observations used for diagnosis.
	Economy of use	Staff commitment to use textbook assessment instruments and tools reflected efforts to expand range of assessment strategies, also saved time.
Method of assessment	Standardized tests	District tests, Special education testing for aptitude and achievement, school-based diagnostic testing (CAT).
	Classroom recitation	Oral questioning common practice for assessment and for learning (scaffolding), interviews, conversations, or group work.
	Performance assessment -spontaneous	Some assessments planned by the teachers, modified from textbook guide, others were part of the instructional process (class activities).
	Performance assessment -formal	Formal performance tasks planned into the mathematics units, selected from the mathematics teacher guide.
	Portfolio collections	Used in writing assessment only in some classes.
	Affective/ social records	Not formalized.
Feedback procedures -oral -non-verbal -written	descriptive vs symbols guidance vs praise immediate vs delayed process vs product oriented fair vs unfair germane vs irrelevant assessment conversation teacher vs student evaluation	Feedback on assessments commonly includes letter grades or percents. Rubric assessments for some junior class also include written comments that focus on strengths and next steps. Achievement levels consistently used for communication and reasoning in journals. • Formative feedback on a regular basis (weekly), focus on process. • Assessments match what was taught, frequently the same format. • Assessments instruments predominantly from the mathematics text and teacher resources. Quizzes and tests may be supplemented by teacher made tests or commercial worksheets linked to curricular expectations. • Criteria used for assessment, but subjective element recognized. Moderation by team partners common practice. • Student self-assessment taught and fostered through use of rubrics and checklists. • Tests and rubrics sent to parents, parental support not consistent or practical due to language issues. Parent-teacher interviews preferred communication.

Figure H 16. Classroom Assessment.

Note. This chart displays data collected and reduced from school interviews, documents and field notes.

Characteristics	Northside School
Goals -acceptance vs resistance	School goals reflect compliance with and acceptance of standards assessment policy, initial resistance directed at the EQAO testing experience changing.
Instruments -collective criteria -portfolio use -exemplars in use	Assessment criteria based on the expectations and outcomes of the curriculum. Students assessed against the levels of performance described in the provincial achievement charts in common practice unless modified by IPRC process (IEP). Common use of district Assessment and Grading Handbook. Some use of portfolios for writing. Provincial exemplars and EQAO anchors used to inform classroom assessment.
Methods (Processes) -planned collection -embedded -toolkit -process goals -diagnostic, formative and summative assessment -assessment literate -collaborative development of tasks and tools	Planning assessments as part of unit development common but not consistent practice. Some development of formative assessment tasks, changes in instructional planning to scaffold learning, meet student needs. Use of embedded assessments a function of the new mathematics textbook program. Common but not consistent practice. Variety of assessment strategies evident but further development of assessments required. Use of formative assessment established, performance tasks beginning to be implemented. Diagnostic, formative and summative assessments are consistent in teacher practice. Feedback provided to students. Assessment literacy developing. Sharing of performance based tasks and assessment tools done formally and informally, best practices being promoted at team meetings, team planning extensive but not consistent.
Personnel -school-wide involvement -student and peer -parent and community	School wide involvement in improvement planning (SIP), commitment to using new assessments. Students involved in self-assessments a common practice. Parent and community participate in School review process and SIP established. Parental involvement in classroom assessment processes encouraged.
Values -teaming valued -appreciates assessment as instruction -uses student assessment -learning skill (process) focus -assessment addresses learning styles -results used for individual student planning -consultation with community	Team approach to assessment common practice at the junior and intermediate levels. Assessment articulated as part of instruction a common practice. Benefits of student self-assessment articulated. Performance assessments common practice, focus on learning skills. Common Practice. Assessment involves a variety of strategies but paper and pencil assessments remain common practice. Use of diagnostic and formative assessment to plan for individual student needs is established practice at junior and intermediate levels. IPRC'd and IEP'd students receive remedial programming. CAT used as diagnostic tool. School staff discusses needs with district staff, parents, Parent Council, other Signal Schools and Community Support groups. Parental support of classroom assessment is limited due to ESL issues, but close communication about student progress established practice and enabled by language liaisons.

Figure H 17. Dynamics of Northside School Assessment Practices.

Note. This chart displays data collected and reduced from district and school interviews, documents and field notes.