



## **CONTENTS**

### **1. Introduction**

### **2. The Meaning Of Accountability**

2.1 Background.....	8
2.2 The meaning of accountability .....	11
2.3 Levels of accountability.....	11
2.4 Who is expected to provide an account? .....	12
2.5 To whom is an account owed? .....	14
2.6 What is to be accounted for? .....	15
2.7 What are the consequences of providing an account? .....	16

### **3. Alternative Approaches to Accountability**

3.1 Alternative perspectives on accountability.....	18
3.2 Market approaches.....	19
3.3 Decentralization approaches.....	21
3.4 Professional approaches .....	23
3.5 Managerial approaches .....	26

### **4. The “Special” Case of Test-based Accountability**

4.1 Why special?.....	29
4.2 Illustrating the complexities .....	30
4.3 Reasons for caution .....	32
4.4 Reasons for action .....	41
4.5 Getting it right .....	42

<b>5. Six Examples Of Whole Accountability Systems</b>	
5.1 Context for accountability .....	47
5.2 Vision and goals .....	48
5.3 Standards .....	49
5.4 Curriculum frameworks and related material .....	49
5.5 Policies: Coherence and integration .....	50
5.6 Performance information .....	51
5.7 Finance and governance .....	52
5.8 Agent receiving information, distributing rewards and sanctions .....	53
<b>6. Toward A Model of Professional Accountability In Saskatchewan</b>	
6.1 Introduction .....	54
6.2 Basic features of a professional approach to accountability .....	55
<b>7. Conclusion</b> .....	61
<b>8. References</b> .....	66

# 1

## **INTRODUCTION**

### **Purposes For The Paper**

Commissioned by the *Saskatchewan School Trustees Association*, this paper was intended to accomplish six purposes. Each of these purposes and how they have been addressed in subsequent sections of the paper are briefly outlined in this introduction.

**1. *Review the current literature and research related to the issue of accountability and transparency in educational governance.***

The paper reviews several different categories of accountability literature. These are literatures concerning the basic concept and meaning of accountability as applied to education and different approaches found around the world for holding schools more publicly accountable. The review examined a more technical literature concerned specifically with the nature and role of testing in most approaches to accountability. Finally, the paper incorporates a growing literature on the actual effects of implementing accountability policies and procedures.

**2. *Describe relevant legislation and regulations in Canadian provinces and other jurisdictions;***

Throughout the paper, different examples of accountability legislation and regulations are mentioned. Accountability policies in most educational jurisdictions, however, are a complex mix of conceptually quite different approaches to accountability. These whole-system approaches to accountability are best reflected in the six examples described in Section 5.

**3. *Outline current popular and best accountability practices related to educational governance in use in Canada and elsewhere.***

The paper touches on most of the accountability policies and tools in use not only in Canada but in other parts of the world, as well. With respect to test-based accountability, in particular, recent developments in British Columbia and Alberta are provided as examples of “getting it right”. The most ambitious approach to accountability at the present time, however, is to be found in England. Tools for accountability in that country are an integral part of nation-wide reform efforts aimed at significantly improving student achievement.

### **4. Address the costs and benefits related to various popular and best accountability practices.**

The paper provides considerably more information about benefits than costs. Furthermore, the evidence reviewed in the paper identifies considerably more negative than positive “benefits”. An overall conclusion from this review, then, is that holding schools more publicly accountable is a relatively weak strategy for school improvement, even though it may be a strategy difficult to resist in today’s political context.

Information about costs provided in the paper lead to two conclusions. First, far too little attention has been devoted to assessing the actual costs of most accountability initiatives. Such inattention may be due to some governments’ assumptions that schools can do more with less. In jurisdictions where this assumption prevails (many U.S. states) no new money is provided for implementing new accountability measures, so costs seem to be constant or reduced. Of course, they are not. At minimum, there are the sort of serious opportunity costs reflected in the narrowing of the curriculum that often occurs in order to enhance test scores.

A second conclusion is that some accountability policies allow for variable costs. The best example of such policies are those concerned with student testing. Clearly it costs more money and more student and teacher time to implement every-pupil testing programs in all grades and all subjects than to implement testing programs focused on several grades, several subjects and based on collecting performance information from only a sample of students or schools.

### **5. Identify considerations that should be had prior to and during implementation of an accountability model both provincially and at a local board level.**

The review of each of four broad approaches to accountability (market, decentralization, professionalization and managerial) includes suggestions for implementing the approach in a manner most likely to result in positive benefits. And Section 4. 5, in particular, outlines a series of lessons learned from previous efforts to implement test-based accountability systems.

### **6. Describe the language of accountability frameworks, including sample language for three or four best models of accountability frameworks suitable for use in education governance tailored for use in Saskatchewan.**

By mutual agreement, this purpose is addressed, in section 6, through the development of a professional, “reciprocal” model of educational

accountability. In this case, reciprocal means that accountabilities are specified for those at all levels of the provincial system, unlike more typical models (and practices) that focus largely on accountability at the school level alone. Such a model acknowledges that the ability of those in schools to achieve the goals of policy makers depends not only on their own practices but the policies and practices of those in many other roles in the provincial educational system.

### Overview

The structure of the paper allows for the six purposes to be addressed both directly and indirectly, throughout, while at the same time authentically reflecting the nature of current conversations underway in both policy and research communities about educational accountability.

Accountability is a much misunderstood concept; so we begin, in Section 2, by clarifying its meaning. Furthermore, efforts to bring conceptual order to the many different approaches to accountability which exist in practice and policy today have produced several competing frameworks; these are compared in the third section of the paper. Guided by one of these frameworks, the largest portion of this section describes four major approaches to accountability and reviews existing empirical evidence about their impact on students, educational professionals and, in some cases, members of the public.

Although not necessary for any of these approaches, as typically implemented all of them rely on student testing for sometimes different purposes; indeed, many people have come to equate accountability with such testing. So Section 4 is dedicated to an exploration of “test-based” accountability, in particular. The next section (5) illustrates, in more detail, the range of currently intact accountability systems using, as examples, policies and practices in six selected Canadian and international jurisdictions. The paper concludes with a proposed model of reciprocal educational accountability for the province.

Terms of reference for this paper linked together accountability and governance. This link seems suitable and the considerable attention to accountability “policies” is the primary way in which governance issues are addressed.

# 2

## **THE MEANING OF ACCOUNTABILITY**

### **2.1 Background**

Greater accountability sometimes has been advocated for schools and school systems simply as a means of demonstrating to taxpayers that they are getting reasonable value for their educational dollar. More often, however, strategies and tools for increasing the accountability of schools are integral parts of much broader school reform initiatives. In many cases those strategies and tools are the central vehicles for reform, on the assumption that holding schools more accountable for what they do and how they do it will itself trigger improvements unspecified and perhaps unimagined by those advocating such accountability. Indeed, it is not an exaggeration to claim that accountability has been the dominant feature of reform efforts in schools since the late 1980's with little sign of diminished interest, as the recent U.S. *No Child Left Behind Act* so forcefully suggests.

The timing of calls for greater accountability on the part of schools and school systems has been remarkably similar across many countries although the reasons for those calls often have not always been the same. In the US, for example, calls for greater educational accountability became quite strident in the late 1960's and early 1970's, subsided briefly only to reemerge with much greater energy in the early 1980's. School choice, a central feature of the U.S. accountability movement during the latter part of this period, arose in a political and social context "unlike any other period in American history" (Cookson, 1994, p.13). During that time the U.S. underwent a cultural transformation – not unlike the change that occurred in Ontario just 8 years ago - in which values supportive of civic ties and obligations were replaced by values associated with personal acquisition and self-interest. "The Reagan revolution challenged the public school system at a very basic level. As the leader of a powerful conservative movement, Reagan legitimated a political philosophy that was essentially antigovernment and probusiness. The Reagan revolution waged ideological war against government spending, wasteful public institutions, and the welfare state in general" (Cookson, 1994, p.6).

“  
*Indeed, it is not an exaggeration to claim that accountability has been the dominant feature of reform efforts in schools since the late 1980's with little sign of diminished interest*  
”

New Zealand provides another instructive example. Concerns giving rise to increased school accountability in that country also began to emerge in the early '60's. These concerns can be traced to dissatisfaction with progress in changing the school system under a highly centralized administrative structure. By the late 1980's discontent on the part of communities and groups that were disadvantaged by the system, such as Maori and women, could not be ignored, leading the government to restructure its education system (as well as health and other systems) to achieve smaller, more effective regional units. A new Education Act came into effect in October, 1989 at which time schools took over their own administration, previous regional education boards were disbanded, and school councils were set up to govern schools (Williams, Harold, Robertson, & Southworth, 1997).

Calls for greater accountability in England, another useful example, grew out of concerns, during the early 1970's over the closed nature of the school curriculum, one largely determined by teachers with little or no influence by parents, business people or government; local educational authorities provided little direct guidance to schools on such matters. In 1976 the chief inspector of schools gave a widely reported speech drawing attention to the problem and the national inspectors' reports began to raise questions about the quality and relevance of what was being taught for the modern world. These and other initiatives eventually culminated in the Education Reform Act in 1998. One part of that Act was concerned with a national curriculum and methods to assess its achievement. The second part of the act focused on the creation of a market system in education based on the philosophy that efficiency and quality are best sustained and enhanced where there is choice and the information needed to make sensible decisions (Williams et al., 1997)

The demands for greater accountability in England were ratcheted up significantly with the government's launching of its National Literacy and National Numeracy Strategies in the late 1990's. The strategies included the establishment of performance targets for individual schools, as well as schools in the aggregate, and schools were (and still are) publicly labeled according to how well they were meeting their targets.

The main source of concern about accountability in Germany was not a diagnosed quality crisis, as in the U.S., but an attempt to cope with dwindling political legitimation and to reduce educational policy conflicts. These conflicts arose in response to the significant social, political, economic, and cultural changes that occurred in the past two and a half decades. A fiscal crisis challenged the welfare state and caused growing awareness of the limitations of the state's problem-solving capacity.



Increased pluralism in the value orientation of society developed in Western Germany in particular and this triggered a fundamental reconsideration of the role of the state, and the nature and scope of the state's function in education. Changes, however, were not driven by a desire to introduce market forces, as in England or New Zealand.

Whereas many tools for increasing school accountability, especially those aimed at increasing competition for students, are of relatively recent origin in many countries, this is not the case everywhere. School choice, for example, emerged as a noticeable movement in the 1980's in the U.S. and some other countries. But the Netherlands has been pursuing its own unique version of school choice for more than 60 years. Since the 1920's parents in that country have been able to choose among Catholic, Protestant and religiously neutral, private sector schools. Furthermore, schools in the Netherlands are financed according to the number of pupils enrolled. So parents dissatisfied with prevailing school options may establish a new school by finding enough parents who will send their children to that school.

With some exceptions, then, the current preoccupation with educational accountability appears to have begun in most developed countries in the 1960's, acquiring significant new energy during the mid-to-late 1980's. The reasons for these calls for greater accountability, furthermore, are to be found in the wider economic, political, and social context of which schools are a part. These contexts are not uniform across all countries. Nevertheless, a core of developed countries including, for example, Australia, New Zealand, the UK, and Canada, have been similarly influenced by New Right and neo-liberal ideologies. Peters has described the context of New Right thinking as follows:

“  
*...reasons for greater accountability, are to be found in the wider economic, political, and social context of which schools are a part.*  
”

A universal administrative reform movement precipitated immediately by a fiscal crisis due to two factors: changes in the international economic system and the spiraling demand for government services. This situation has been variously described as the 'crisis of the welfare state,' as a breakdown of the Keynesian post-war consensus, and as leading to a global shift to the Right in western capitalist democracies. In general terms, policy responses have aimed first and foremost at greater budgetary restraint pursued in a variety of ways: the downsizing of public-funded organizations; corporatization and privatizing strategies; deregulation; and asset sales programmes (1992, p.269).

## 2.2 The Meaning of Accountability

The dictionary meaning (e.g., Webster's 7th New Collegiate Dictionary) of accountability and its near relatives suggests that to be accountable means being subject to giving an account, being answerable, and capable of being accounted for. The term "account" entails giving a report on, furnishing a justifying analysis or explanation, providing a statement of explanation of one's conduct, offering a statement or exposition of reasons, causes, grounds, or motives, or simply providing a statement of facts or events<sup>1</sup>. Beyond this general definition of the term, the specific nature of an accountability policy, mechanism or tool depends on the answers to five questions.<sup>2</sup>

- What level of accountability is called for?
- Who is expected to provide the account?
- To whom is the account owed?
- What is to be accounted for?
- What are the consequences of providing an account?

“*The term “account” entails giving a report on, furnishing a justifying analysis or explanation, providing a statement of explanation of one’s conduct, offering a statement or exposition of reasons, causes, grounds, or motives, or simply providing a statement of facts or events<sup>1</sup>*”

## 2.3. Levels of Accountability

As the dictionary definitions of accountability alluded to above begin to clarify one of the distinctions central to a more fundamental conception of accountability is the level of accounting required. For example, included within the meaning of “providing an account”, according to dictionary definitions, is the relatively simple description of events. A school which distributes a calendar outlining its program options for students along with other events, policies and the like, that it considers noteworthy for students and parents is “providing an account” in the *descriptive* meaning of the term. This is the simplest or lowest level of accountability.

Also referred to as part of the meaning of account, however, is the more difficult offering of an *explanation* for events. The typical school calendar would not suffice as an account given this meaning of the term, unless added to the calendar was further information about the sources of programs, the reasons for other events, and the like. Such a calendar, however, would still fall short of “providing an account” when the term is defined to include

---

<sup>1</sup> From a systems perspective, there is no reason to excuse any level of the provincial education system from the need to provide an account of their actions.

<sup>2</sup> See Wagner (1989)

*justification* for events, the highest level of accountability. This meaning of “providing an account” would require the calendar to contain not only descriptions and explanations. It also would require arguments of some sort for why these programs and events were the most appropriate ones for the school to be offering its students.

### 2.4 Who is expected to provide an account?

“Responsibility” is one of two minimum conditions for validating the assignment of any accountability obligation and whatever interpersonal relationships it may entail, according to Wagner’s (1989) analysis. The assignment of any responsibility or obligation to provide an account is usually depends on identifying the person or parties responsible for the act creating it. This depends on the act itself and the institution within which the act is carried out. So, one becomes obliged or responsible to provide an account as a result of either an act that one undertakes or the role that one occupies within an organization.

“*The assignment of any responsibility or obligation to provide an account is usually depends on identifying the person or parties responsible for the act creating it.*”

A person or group can be causally responsible for an act which is self evident and ultimately personally obliging. This may be an act actually committed by the person or group, as in the obligation a teacher assumes for the welfare of her students on a field trip. Obligations also arise when the person or group exercises influence on others to perform an act. This type of obligation is acquired by a principal, for example, in relation to the quality of instruction in a teachers’ classroom, when that principal encourages the teacher to implement innovative curriculum material.

Persons or groups may be responsible, as well, for the omission of acts - acts in which the person or groups engagement is considered inappropriate. The teacher taking students on a field trip to the zoo is expected not to detour through the local shopping mall on the way to the zoo. Such persons or group also may be causally responsible for an act which is initially but not ultimately obligating because of mitigating circumstances. While the teacher is initially responsible for the safety of his students on a field trip, various natural disasters (e.g., lightening or other “acts of God” as the insurance industry calls them) would be considered beyond the teachers control, thus absolving the teacher from responsibility, providing that he took appropriate precautions on behalf of his students.

And joint responsibility may be acquired for an act incurring a shared obligation among those responsible. The statutes and laws governing

elementary and secondary education in most provinces and countries, for example, hold teachers and principals jointly responsible for students' educational experiences. Individual or joint responsibility may be acquired for an act obligating others, as well. In some jurisdictions (e.g., Ontario) principals and superintendents are jointly responsible for the quality of instruction that teachers provide their students.

One acquires obligations not only through the acts one carries out oneself (creating causal obligations) but also throughout the roles or positions one holds in life: for example, educators historically have been expected to behave in both their personal and professional lives in ways that are consistent with the norms of morality held by the communities in which they work. Such roles create non-causal or "expectational" obligations of either a specific (e.g., what we do as teachers), or general nature ( what we are expected to do as good neighbors or citizens). Expectational obligations entail potential acts or performances that are possible to fulfill, consistent with the role from which they are said to derive, and reasonable to expect in light of other considerations that are equally relevant. Most people have acquired both causal and expectational types of obligations and these types become very difficult to distinguish in real-life circumstances.

While the distinctions around different forms of accountability outlined above, may seem to be a bit esoteric or overdrawn on first encounter, they are, in fact, highly germane to the framing of accountability policies and practices for schools. To illustrate, it is highly questionable, following Wagner's (1989) reasoning, whether a person should be held accountable for acts which, causally, she has neither omitted, committed or influenced; some will argue that ensuring students come to school properly fed and ready to learn is an example of this. It is equally questionable whether a person should be held accountable for expected performances which are impossible to satisfy (e.g., ensuring that all students learn to high standards), are inconsistent with the role from which they are said to derive (e.g., teachers being held responsible for students' use of illegal substances), or whose assignment and satisfaction may be quite unjustified by other factors (principals being held responsible for improving the average reading scores of students in schools with highly transient student populations).

“

*Also it is questionable whether a person or an organization should be held solely accountable for matters involving a shared, causal responsibility.*

”

Also it is questionable whether a person or an organization should be held solely accountable for matters involving a shared, causal responsibility. The

success of students in school, for example, is a function of many factors. While the quality of teachers' instruction is important, it is significantly influenced by such factors over which the community or the government, not the teacher or the school, has control such as the physical condition of the school building, the size of classes, the time available for teachers to prepare for classes, and the like.

Nor is it legitimate to hold a person solely accountable for expected performances requiring a shared influence, unless it is specifically understood that the person is answerable for the actions of others, as would seem to true of parents' responsibility for the actions of their children, or executives' responsibility for the work of subordinates. When we know, for example, that family educational culture accounts for at least 50% of the variation in student achievement, how is it possible to claim that schools alone should be held accountable for such achievement?<sup>3</sup>

Somewhat grudging acknowledgment of the complex nature of responsibilities for the many factors that influence the learning of students has led to recent arguments for developing "reciprocal" accountability systems, the features of which are described later in this report.

### 2.5 To Whom Is An Account Owed?

"Entitlement" of the person or group requesting an account is the second minimum condition for justifying such a request. Entitlement is a function of whether a legitimate interest can be shown by those expecting an account. Such an interest must be demonstrated in order to validate an obligation to satisfy an accountability demand.

There are different degrees of entitlement to an account and one's entitlement increases with the degree of valid interest one has in the act for which the account is requested. This is sometimes quite difficult to determine. For example, when a teacher is required to provide an account of

“  
*Somewhat grudging acknowledgment of the complex nature of responsibilities for the many factors that influence the learning of students has led to recent arguments for developing “reciprocal” accountability systems.*

”

“  
*Entitlement is a function of whether a legitimate interest can be shown by those expecting an account.*

”

---

<sup>3</sup> This common estimate is from research which attempts to identify what classroom, school and other factors explain differences in students' achievement across schools. Initial attention was drawn to the importance of family background characteristics by Coleman (1966) and his colleagues. It should be noted that there is often more variation in student achievement (as well as educational practices) within than across schools.

their classroom instruction through participation in a performance appraisal process, it seems clear that the appraisee and the appraiser are entitled to the account. But who else might have a legitimate entitlement to the account? Typically, very few others have access to the results of the appraisal (perhaps a senior school system official through the personnel files kept in the school system's office). But what is the entitlement of the parents of the teacher's students, and the students themselves? It does not seem difficult to justify the pre-eminence of their entitlement, although many reasons unrelated to entitlement have been developed to keep formal accounts of teacher performance out of reach of those whose entitlements may be greatest.

Instead, the closest parents and students typically get to a formal account of teacher performance is a report of student achievement and such reports usually are considered to be more relevant to student than teacher accountability. Indeed, this is clearly a flawed basis for teacher accountability since teachers, parents, students, peers, and a variety of physical and social conditions combine in their contribution to student achievement. Responsibility for student achievement is jointly shared, whereas responsibility for teacher performance is much more unambiguously teachers' responsibility, albeit one shared in some degree by those who help create the conditions of teachers' work.

Teachers, administrators, parents and policy makers will sometimes hold quite different opinions about who is "really" entitled to an account. Evidence collected from Ontario teachers for example, indicated teachers' strong feelings of accountability to their students and parents, moderate accountability to school administrators and almost no sense of accountability to the provincial government (Leithwood, Jantzi, & Steinbach, 2002). Similar findings have been reported recently from teachers in both Maryland and Kentucky (Mintrop, 2004). Teachers in both studies were willing to acknowledge their legal obligations for some distant form of accountability to the government but did not consider it to be especially meaningful.

“  
*Teachers, administrators, parents and policy makers will sometimes hold quite different opinions about who is “really” entitled to an account.*  
”

### 2.6 What is to be accounted for?

In an educational system, what is to be accounted for most fundamentally is the welfare of individual students. And within the range of possible meanings of “welfare”, most educational systems these days focus on academic achievement, often the results of system-wide tests. Such achievement is a necessary if not sufficient part of the meaning of student welfare in the context of schooling.

In current policy and practice, educators are often held accountable, as well, for organizational conditions and practices believed to contribute more or less directly to students' welfare. One prominent example is characteristics or conditions of schools explaining variation in their effectiveness (e.g., collaborative professional cultures, high expectations for student achievement, clear goals). These organizational conditions figure among the criteria government inspectors use in evaluating schools in Scotland and England, for example.

Another example of what is to be accounted for, other than student welfare directly, are standards of professional teaching practice (elaborated in more detail below). In the U.S. context, such standards for advanced teaching have been developed by several agencies, among them the National Board for Professional Teaching Standards (1998). Parallel standards for beginning teaching have been developed by the Interstate New Teacher Assessment and Support Consortium (1992). The Center for Research on Educational Accountability and Teacher Evaluation developed a taxonomy of teaching "duties" (Scriven, 1994), something else for which educators may be held accountable.

As these illustrations begin to suggest then, in addition to being held accountable for student welfare directly, educators in some contexts are held accountable for:

- ensuring that specific organizational qualities or conditions considered to be critical to effectiveness are reflected in their schools or districts;
- organizational efficiency
- meeting standards of professional knowledge and skill;
- meeting standards of moral behavior; and
- implementation of best professional practices or specified duties.

### 2.7 What are the consequences of providing an account?

Providing an account may trigger three responses on the part of the receiving person or group. The first and least consequential of these responses occurs when accounting is voluntary. This would be the case when a teacher voluntarily sends out a newsletter to parents describing what their students program will be like in the next month. In a case such as this,

“  
*In current policy and practice, educators are often held accountable, as well, for organizational conditions and practices believed to contribute more or less directly to students' welfare.*  
”

the voluntary nature of the account reduces the likelihood of any response at all that could have connotations of accountability.

A second type of consequence is most likely to occur when an account is obligatory but no consequences have been formally specified. In such cases it seems likely that some response will occur but this response often will be muted and, almost by definition, unpredictable. The requirement, in the Canadian province of Alberta, that school districts annually publish district profiles is an example of an obligatory account unlikely to provoke a predictable response because the form of the response is unspecified. In this case the obligation is a legal one, an obligation spelled out in policy.

But an account also may be considered obligatory on moral grounds. That is, the person or group providing the account may feel that the actions for which they are responsible carry with them an obligation to account by virtue of the special nature of the responsibility. It might well be case, for example, that the teacher's newsletter is stimulated by a sense of moral obligation the teacher feels to report to parents about the upcoming experiences of the students which their parents have entrusted to her care.

Finally, there are circumstances where an account is required, and rewards and punishments for the person(s) providing the account are specified. Circumstances such as this are now increasingly common in many jurisdictions. In England, schools judged to be failing are publicly identified and provided with a team of external educators to turn things around. In some U.S. states schools that fail to meet achievement targets over specified periods of time may be placed under review, principals and teachers may be required to implement specific improvement measures, be reassigned to other schools and/or the school may be "reconstituted". Florida currently has among the most draconian of such policies, with families receiving tuition to move their children to another public or private school if they find themselves in a "failing" school.

Wagner (1989) argues that in cases where there is no requirement or obligation that an account be given there is no accountability. But even when only very simple types of reports are expected (descriptions), if there is a requirement or obligation, then this becomes accountability.

So some form of obligation or requirement is an essential part of accountability from his perspective. This position rules out, as a form of accountability, the teacher's non-obligatory monthly newsletter, but includes the provincially - required annual report and meeting with ratepayers on the part of school boards.



# 3

## **ALTERNATIVE APPROACHES TO ACCOUNTABILITY**

In Section 2, the meaning of accountability was clarified in response to five questions: who is accountable, at what level, to whom, for what, and with what consequences. This section describes different approaches to accountability. We begin by comparing and contrasting two different efforts to bring order to the wide variety of accountability mechanisms evident in education systems at this time. Using one of these frameworks, we then describe four quite different approaches to accountability; the policy mechanisms or tools associated with each of these approaches are described and evidence about their effects is summarized.

### **3.1 Alternative Perspectives On Accountability**

Two category systems have been created to help better understand the similarities and differences among the many different mechanisms now used to increase the accountability of public schooling. Developed exclusively in a U.S. context, the first of these systems (Adams & Kirst, 1999; Henig, 1994) includes six categories – bureaucratic, legal, professional, political, moral and market approaches to accountability. Arising from analyses of international accountability initiatives, the second system (Leithwood, 2001; Leithwood & Earl, 2000), includes four categories – market, decentralization, professional and managerial approaches.

The two systems share several key similarities. Both include two of the same categories (market and professional forms of accountability). While moral and legal concerns are not explicit categories in the second system, as they are in the first, the second system does include treatment of moral and legal issues within its categories. And issues addressed in the bureaucratic category of the first system are subsumed in the decentralization and managerial categories of the second system.

But there are important differences: unlike the second system, a key aim of the first is to clarify principal/agent relations - relations between those who are held accountable and those to whom an account is owed; unlike the first, more inclusive, system, the second focuses mostly on the relatively

“  
*In practice, most government accountability policies are a “mixed bag” including policy levers or mechanisms from several or all approaches.*  
”

recent performance or outcome (student learning) oriented mechanisms for accountability, as distinct from mechanisms concerned also with inputs (e.g., funding) and processes (e.g., school organization) for schooling. This second, four-category system is used as the framework for the remainder of this section because of its performance-based orientation.

All four approaches to accountability described in this section assume that increasing accountability will improve the quality, and in two cases, the efficiency of education that schools provide their students; student learning will improve, as a consequence. But each approach is based on a different, and sometimes conflicting, set of assumptions about what is wrong with schools and how to fix them; so each employs a largely different set of policy mechanisms or tools, although student testing, addressed in a subsequent section of its own, can be found in the tool chest of all four approaches. It needs to be said that this classification system is a conceptual and analytic convenience. In practice, most government accountability policies are a “mixed bag” including policy levers or mechanisms from several or all approaches (a powerful explanation for the confusion such policies often create for those having to implement them).

### **3.2 Approaches to Accountability**

<b>Market Approaches</b>	<b>Decentraization Approaches</b>	<b>Professional Approaches</b>	<b>Management Approaches</b>

Sometimes referred to as the exit option, market approaches to accountability increase competition among schools for students. An especially prominent approach currently, versions of it are evident in several European countries, Canada, the United States, New Zealand, Australia and parts of Asia, for example. Specific tools for increasing competition among schools for student-clients include allowing school choice by opening boundaries within and across school systems, school privatization plans, the creation of charter schools, magnet schools, academies and other specialized educational facilities. Competition also is increased by altering the basis for school funding so that the money follows students (e.g., vouchers, tuition tax credits), and by publicly ranking schools based on aggregated student achievement scores. These tools are often used in combination.

“  
*...market approaches to accountability increase competition among schools for students.*  
”

Among advocates of these different tools is a common belief that schools are unresponsive, bureaucratic, and monopolistic (Lee, 1993). Members of such organizations are assumed to have little need to be responsive to pressure from their clients because they believe they are not likely to lose them. In relation to schools, this means that they will come to view their major task as offering programs that *they* believe are good for their clients. Members of such organizations, it is argued, seek efficiency on their own terms and are prone to view clients as objects “to be treated” rather than customers “to be served”.

Advocates of market approaches to accountability (e.g., Chubb & Moe, 1990) hold a series of assumptions about how such competition is likely to result in greater student achievement. First, increased competition allows parents and students to select schools with which they are more satisfied and which better meet their educational needs. Second, parents who are more satisfied with their child’s school provide greater support to that school and to their child’s learning. Third, students are likely to be more engaged when their own learning styles are matched to a particular school. Fourth, when teachers have chosen their work settings and have been active in designing their own schools’ programs, they will be more committed to implementing those programs effectively. Finally, all of these outcomes will combine to increase student achievement, attendance, and educational attainment (Elmore, 1990; Raywid, 1992).

*Evidence of impact and suggestions for implementation.* What about evidence concerning the actual effects of markets on the behavior of school people? This evidence suggests, first, that choice arrangements vary considerably in the autonomy awarded those at the school level. As an explanation for the few differences found in the practices of U.S. principals of magnet and nonmagnet schools, Hausman (2000) pointed to the wide array of district policies regulating all principals in the district. Second, evidence demonstrates that some school choice settings actually put very little pressure on schools to compete. This is the case when a school is oversubscribed (Hausman, 2000), or when it serves parents and students who, for economic and other reasons, feel unable to travel to a school outside their own neighborhood (Lauder & Hughes, 1999).

Finally, schools facing the same competitive conditions may respond quite differently for reasons associated with the individual abilities, values, beliefs, and motivations of their administrators and teachers. For example, Grace’s study (2002) identified three quite different responses by individual school leaders to increased competition for students. One group

of leaders welcomed the more managerial role they believed was implied in policy changes. A second group were preoccupied with the loss of a professional orientation to schools, and concerned about managerialist values encroaching on their work. The third group of school leaders actively opposed those features of market approaches to school reform which they believed were unlikely to lead to school improvement. Other evidence suggests that competition has unpredictable effects on the propensity of school leaders to engage in instructional leadership, some finding little time for it while others increase their attention to it (Hausman, 2000).

This brief review of theory and evidence suggests, most obviously, that school staffs implementing market solutions in truly competitive environments need marketing and entrepreneurial skills. By themselves, however, such skills do not acknowledge the growing evidence that market approaches to accountability can be, and usually are, highly inequitable (Lauder & Hughes, 1999; Lee, 1993). When equity is a strongly valued goal, schools will need the ability to market themselves in ways that make access possible even for those children and families from diverse and economically disadvantaged backgrounds (Bauch & Goldring, 1995).

### **3.3 Approaches to Accountability**

<b>Market Approaches</b>	<b>Decentraization Approaches</b>	<b>Professional Approaches</b>	<b>Management Approaches</b>

When decentralization of decision making is used for purposes of increasing accountability, one of its central aims often is to increase the voice of those who are not heard, or at least not much listened to, in the context of typical school governance structures. When this is the goal, a *community control* form of site-based management (e.g., Wohlstetter & Mohrman, 1993) typically is the instrument used for its achievement. The basic assumption giving rise to this form of site-based management is that the curriculum of the school ought to directly reflect the values and preferences of parents and the local community (Ornstein, 1983). School professionals, it is claimed, typically are not as responsive to such local values and preferences as they ought to be. Their responsiveness is greatly increased, however, when the power to make decisions about curriculum, budget, and personnel is in the hands of the parent/community constituents of the school. School councils in which parent/community constituents have a majority of the membership are the primary vehicle through which to exercise such power.

Devolution of decision making, however, is sometimes rooted in a broader reform strategy for public institutions, which Peters has referred to as “new managerialism.” According to Peters, new managerialism “...emphasizes decentralization, deregulation and delegation” (1992, p. 269). While there are variants on this approach to accountability among countries, they share in common a shift in emphasis (a) from policy formulation to management and institutional design; (b) from process to output controls; (c) from organizational integration to differentiation; and (d) from ‘statism to subsidiarity ’ (Peters, 1992).

In countries such as New Zealand and Australia where school reform has been substantially influenced by the philosophy of new managerialism, creating more efficient and cost effective school administrative structures is a second central goal for devolution. Typically, this goal is pursued through the implementation of an *administrative control* form of site-based management which increases school-site administrators’ accountability to the central district or board office for the efficient expenditure of resources. These efficiencies are to be realized by giving local school administrators authority over such key decision areas as budget, physical plant, personnel, and curriculum. Advocates of this form of site-based management reason that such authority, in combination with the incentive to make the best use of resources, ought to get more of the resources of the school into the direct service of students. To assist in accomplishing that objective, the principal may consult informally with teachers, parents, students or community representatives. Site councils are typically established to advise the principal but with membership at the discretion of the principal.

Decentralization approaches to accountability assume that school leaders will become teachers of those with newly found voices, usually parents and/or staff. The school leader’s task is to “empower” these people and to actively encourage the sharing of power formerly exercised by the principal (Tanner & Stone, 1998). School leaders, it is assumed, will act as members of teams rather than sole decision makers. This role entails helping others to make defensible decisions and clarifying their decision responsibilities. As well, school leaders will embrace the belief that, through participation in decision making, teachers and parents will not only be more committed to the results of such decision making, but that the decisions themselves

“  
*The basic assumption giving rise to this form of site-based management is that the curriculum of the school ought to directly reflect the values and preferences of parents and the local community*

“  
*School councils in which parent/community constituents have a majority of the membership are the primary vehicle through which to exercise such power.*

will be better. The school leader becomes the keeper of the process, not the outcome of the process (Harrison, Killion, & Mitchell, 1989; Williams et al., 1997).

*Evidence of impact and suggestions for implementation.*

Evidence of the effects on school organizations of decentralization or school based management in its various forms is quite extensive (Bullock & Thomas, 1997; Leithwood & Menzies, 1998, Tanner & Stone, 1998; Wildy & Loudon, 2000). This evidence indicates that while sometimes the advantages of decentralization are realized, often they are not and other less positive outcomes occur. Decentralization is associated, for example, with a radically increased emphasis for administrators on budgetary considerations and less attention to providing leadership about curriculum and instruction (Daresh, 1998). Decentralization greatly increases the time demands on school staffs, especially administrators, and the need for more attention to time management (Cranston, 2000), intensifies their role (Williams, Robertson & Southworth, 1997) and, in quasi-market conditions, may isolate them from other administrative colleagues outside their own organization.

In terms of impact on teaching and learning, considerable empirical evidence suggests quite disappointing effects (Leithwood & Menzies, 1998). In those exceptional cases where teaching and learning have benefited from this approach to accountability, school leaders have, for example, adopted a supportive leadership role themselves, nurtured leadership on the part of others, and strongly encouraged councils to adopt a capacity-building agenda (Beck & Murphy, 1998). Leadership practices such as these help transform an otherwise impotent strategy into at least a modest force for improving teaching and learning.

“  
*School leaders, it is assumed, will act as members of teams rather than sole decision makers.*  
 ”

**3.4 Approaches to Accountability**

Market Approaches	Decentraization Approaches	Professional Approaches	Management Approaches

There are two radically different accountability strategies that have a professional orientation. One of these approaches manifests itself most obviously in the implementation of *professional control* models of site-based management. The other approach encompasses the professional standards movement as it applies to the practices of teachers and

administrators. What proponents of both strategies share in common is a belief in the central contribution of professional expertise to the outcomes of schooling. They differ most obviously on which practices they choose for their direct focus. In the case of professional control site-based management, the focus is on school-level decision making, whereas classroom instructional practices and school leadership practices are the primary focus of the professional standards movement.

Professional control site-based management (Murphy & Beck, 1995) increases the power of teachers in school decision making while also holding teachers more directly accountable for the school's effects on students. The goal of this form of site-based management is to make better use of teachers' knowledge in such key decision areas as budget, curriculum and, occasionally, personnel. Basic to this form of site-based management is the assumption that professionals closest to the student have the most relevant knowledge for making such decisions (Hess, 1991), and that full participation in the decision-making process will increase their commitment to implementing whatever decisions are made. Participatory democracy, allowing employees greater decision-making power, is also presumed to lead to greater efficiency, effectiveness and better outcomes (Clune & Witte, 1988). School councils associated with this form of site-based management typically have decision-making power and, while many groups are often represented, teachers have the largest proportion of members.

A standards approach to accountability in the traditional profession emphasizes heavy control of entry to the profession by government with responsibility for subsequent monitoring of accountability turned over to members of the profession itself (e.g., colleges of physicians, lawyers' bar associations). Such an approach requires clear standards of professional knowledge, skill, and performance, something the professional standards movement in education set out to define beginning in the U.S., for example, in the early 1980's. Different products of the standards movement are available as the basis for the licensure of entry-level teachers (e.g., INTASC's Model Standards for Beginning Teacher Licensing, Assessment and Development) and school administrators (e.g., State of Connecticut Department of Education) as well as for recognizing advanced levels of teaching (e.g., The National Policy Board for Teaching Standards), and school administrator performance such as the Interstate School Leadership Licensure Consortium standards (ISLLC) and Education Queensland's "Standards Framework for Leaders").

“  
...both strategies share  
in common is a belief  
in the central  
contribution of  
professional expertise  
to the outcomes  
of schooling.”

“  
A standards  
approach to  
accountability in  
the traditional  
professions  
emphasizes heavy  
control of entry to  
the profession by  
government”

Wise and Leibbrand (2003) have developed a model of professional accountability focused especially on teachers. This model, while heavily U.S. centric, is provocative nonetheless. It builds on compelling evidence, reviewed by the authors, that student success at school is significantly dependent on the level of teachers' content and pedagogical knowledge and skill. According to this model, teachers are accountable for implementing those best practices spelled out in explicit statements of teaching standards. However, unlike many current approaches to holding teachers accountable, this model explicitly acknowledges that what teachers do is fostered and constrained by the actions of many others. So the model holds a wide range of roles and institutions accountable for ensuring that teachers are capable of implementing best practices. These accountabilities, or actions to be taken, are specified for governments, faculties of education, districts (local boards of education are to be accountable for formulating teacher compensation to reward accomplished teaching and support tiered licensing) and school administrators. Student and parent accountabilities are also included in this model. As this comprehensive range of accountabilities suggests, the model is based on a staged conception of teacher capacity development, beginning with recruitment and extending through initial preparation, induction, early mentoring to the development of mature expertise.

“  
*Professional approaches to accountability imply an increased need for school leaders to stay abreast of best professional practices and to assist staff in the identification of professional standards for their work.*  
”

Professional approaches to accountability imply an increased need for school leaders to stay abreast of best professional practices and to assist staff in the identification of professional standards for their work. School leaders, in the context of professional approaches to accountability, need to both set expectations and create conditions for professional growth (Prestine, 1999). Also, these leaders need to: monitor progress of staff toward the achievement of professional standards; buffer staff from external distractions; assist parents to understand and appreciate such standards; and mobilize resources to meet not just higher but more sophisticated standards. Maintaining teacher morale in schools identified as low-achieving, and helping ensure equitable treatments for the needs of students also is a challenge for school leaders responding to this form of accountability (Bay & Reys, 1999)

“  
*To ensure that the content of whatever improves student learning are reflected in practice.*  
”

*Evidence of effects and suggestions for implementation.* While a substantial amount of evidence suggests that site-based management, in general, contributes little to student learning, teacher-controlled forms of site-based management are a modest exception (Leithwood & Menzies, 1998); of all



forms, teacher dominated school councils are the only type at least weakly associated with improvements in teaching and learning.

There is little direct empirical evidence about the effects of professional standards on student achievement and what evidence there is provides no clear answer. This lack of evidence, however, should not be interpreted to mean that professional standards have no role to play in improving schools. There is a well-established relationship between teacher knowledge and skill and student learning (e.g., Darling-Hammond & Youngs, 2002). So the challenge for policy makers considering the use of professional standards is twofold:

- To ensure that the content of whatever standards are developed or selected adequately captures the professional knowledge and skill that demonstrably improves student learning:
- To nest the standards within other policies and procedures concerning recruitment, initial training, induction, evaluation and ongoing professional development in order to ensure that the knowledge and skills encompassed by the standards are reflected in practice.

**3.5 Approaches to Accountability**

Market Approaches	Decentraization Approaches	Professional Approaches	Management Approaches

This approach includes systematic efforts to create more goal-oriented, efficient, and effective schools by introducing more rational procedures for doing business. The main assumption underlying this approach is that there is nothing fundamentally wrong with current school structures. The effectiveness and efficiency of schools will be improved, however, as they become more strategic in their choices of goals, and more planful and data-driven about the means used to accomplish those goals.

This approach encompasses a variety of procedures for “strategic planning”, especially at the LEA or district level, as well as multiple procedures for school improvement planning (see the states of Illinois, Florida, and Missouri, for example), school development planning (Giles, 1997), and monitoring progress (e.g., the accountability reviews managed by New Zealand’s Education Review Office).

“*This approach includes systematic efforts to create more goal-oriented, efficient, and effective schools*”

Strategic planning, a highly rational form of planning, has been especially popular at the district level and typically includes a fairly common set of steps. For example, the first half dozen steps in Cook's 11 step model include:

- Determine the organizations fundamental beliefs and values;
- Describe the organization's mission;
- Identify strategic policies for accomplishing the mission;
- Identify the strengths and weaknesses of the organization related to accomplishing its mission;
- Devise a set of structures to enable the mission to be accomplished;
- Identify other organizations (competitors) with the same mission.

School improvement planning typically includes such steps as: examining evidence of strengths and weaknesses, especially in respect to student achievement, setting a small number of goals, developing strategies to accomplish those goals, assigning responsibilities, and monitoring progress.

*Evidence of impact and suggestions for implementation.* In their recent, extensive, review of evidence, Beach and Lindahl (2004) paint a skeptical picture of the value of strategic planning for district and school improvement. They conclude that this form of planning fails to acknowledge, for example, that educational decision makers rarely have available anything like all of the information that would be required for a fully rational decision. Nor are schools actually free to chose their missions. Their missions are mostly chosen for them (by provincial governments). Furthermore, there are close limits on the flexibility with which schools and districts can apply their resources. For those wishing to implement some form of district-wide planning, then, Beach and Lindahl (2004) recommend (a) careful evaluation of the planning process and its contribution to student learning, and (b) consideration of less rational, more incremental approaches to planning, several of which are described in their paper.

School improvement planning (SIP) is at least as popular, at the school level, as strategic planning is at the district level. Nonetheless, there is no more evidence of its value in improving student achievement. As Leithwood et al conclude from a recent review of such evidence, along with the results of their own study:

...by allowing SIP to be the organizing concept for improvement we perpetuate a linear, mechanistic and superficial way of thinking about the improvement problem, whereas powerful solutions are more likely to be non-linear, indirect and embedded deeply in the "guts" of a

school's anatomy – its culture, its reward structures, the dispositions and motivations of those attracted to it as a workplace, and the basic “technologies” it uses to accomplish its goals (in press).

In spite of the typically weak direct effects of planning on student achievement, both strategic and school improvement planning processes bring with them other benefits which may be hard to resist. Planning and plans allow administrators and policy makers to exercise some degree of control over an otherwise autonomous and independent-minded group of educational professionals. Planning process also can generate a considerable amount of new learning among participants about what their organization does, its potential for the future and both its strengths and weaknesses. These advantages may well carry the day, even though planning's contribution to the “bottom line” remains hard to document.

Planning that achieves benefits of the sort outlined here assume that those in leadership roles are skilled in collecting and interpreting systematically collected data and that they are able to develop with their staffs clear, manageable, goals and priorities for school improvement. Leaders are also, it is assumed, able to monitor and refine plans accordingly. Because district resources and cooperation often are needed to accomplish school priorities, school leaders also must be able to develop especially good working relations with their district colleagues. Ensuring such capacities is an important aspect of implementing improvement planning at either the district or school level.

Results of two projects carried out by the University of Cambridge Institute of Education (reviewed by Southworth, 1998) suggest that successful school improvement depends on establishing and sustaining a culture of inquiry and reflection, a commitment to collaborative planning and staff development, high levels of stakeholder involvement, and effective coordination strategies. Establishing these conditions depended on school leaders emphasizing the use of systematic evidence, focusing on student learning, and encouraging careful monitoring of both teaching and pupil progress. “Strategic management” in these projects also entailed developing school improvement plans from the results of inquiry and reflection, and carefully monitoring and evaluating the implementation of those plans.

“*Strategic management*”  
*entailed developing school improvement plans from the results of inquiry and reflection, and carefully monitoring and evaluating the implementation of those plans.*”

# 4

## THE “SPECIAL” CASE OF TEST-BASED ACCOUNTABILITY

### 4.1 Why Special?

Test-based accountability is labelled “special” here because large-scale student testing, in one form or another, is a component or tool included in all four approaches to accountability described in Section 3. Each approach rationalizes the need for test results on somewhat different grounds. Market approaches, for example, conceive of test results as a primary tool for parents to use in choosing the best school for their children, whereas Managerial approaches imagine using test results to help decide on the goals for school improvement plans.

A considerable literature on accountability, in addition, treats test-based accountability as an independent approach in its own right, the labels “test-based”, “performance-based” and “standards-based” being used interchangeably to qualify the meaning of accountability. When this is the case, a small number of other features typically frame the nature and use of tests. For example, Newman, King and Rigdon argue that a complete performance-based accountability system should include at least four parts:

1. Information about the organization’s performance (i.e., test scores).
2. Standards for judging the quality or degree of success or organizational performance (e.g., a mean achievement score higher than other schools with comparable demographic characteristics).
3. Significant consequences to the organization [...] for its success or failure in meeting specified standards.
4. An agent or constituency that receives information on organizational performance, judges the extent to which standards have been met, and distributes rewards and sanctions ... (1997, p.36)

No matter the overall approach to accountability, however, large-scale testing is typically carried out, and the results used, in very similar ways. This section is a “primer” on the issues with which trustees should be familiar as they begin to frame recommendations on student testing for

“*Test-based accountability is labelled “special” here because large-scale student testing, in one form or another, is a component or tool included in all four approaches to accountability*”

provincial consideration and to clarify the purposes and procedures for student testing as part of their own districts' more comprehensive approaches to accountability.

Kornhaber explains that the beliefs giving rise to testing as a strategy for holding schools accountable and improving their contribution to student learning is based on:

...an often unarticulated, logical and rather behaviorist theory...[Students'] performance is suboptimal partly because standards are low and poorly articulated. Therefore, higher standards must be clearly and publicly spelled out. To find out whether students and educators are focusing on reaching these standards, students will be tested. To make sure the new standards and tests motivate teacher and student effort, test results will carry consequences or stakes....To avoid punishments and get rewards, students and teachers will work harder and as a result, students will learn more and be better prepared for the workforce" (2004, pp. 51-52).

As Newman, King and Rigdon (1997) point out, however, this theory is much better suited to commercial organizations than it is to the world of schools. So the effects of test-based accountability are often not what its advocates intend. In a commercial enterprise, for example: "High customer satisfaction leads to increased demand for a product or service and higher profits, which the producer will act to maximize. Customer dissatisfaction leads to lower demand and economic loss..." (p. 36).

### **4.2 Illustrating The Complexities**

A brief story from Ontario will help sensitize you to some of the major potholes into which you may disappear should you chose to argue for a more extensive test-based accountability system in Saskatchewan; similar stories from other countries, provinces and states are not difficult to find.

Over the past three years, "only" about 55% of Ontario students taking the Educational Quality and Accountability Office's (EQAO) grades 3 and 6 reading, writing, and math tests have achieved level 3 performance, the officially sanctioned "acceptable" performance level (on a scale of 1 to 4) for students in the province. These results are widely reported by the media to be an indication of serious weaknesses in the province's education system – "Almost half the kids are failing!" - an interpretation widely accepted by members of the public. Indeed, responding to public angst on this matter, in the spring of 2004 the new Liberal government committed

itself to major initiatives in literacy and math aimed at increasing the proportion of students achieving at level 3 from about 55% to 75% within the four year period of its mandate.

This widely shared view of “the problem” and how to fix it, notwithstanding, results of IEA’s *Progress in International Reading Literacy Study* (PIRLS), conducted with grade 4 students only three years earlier in 36 countries, indicated that only Swedish students scored higher than Ontario students. Similarly OECD’s *Program for International Student Assessment* (PISA), measuring the reading literacy of 15 year olds in 32 countries in 2000, indicated that the mean reading levels of students in Canada, as a whole, and Ontario in particular, ranked second - just below Finland; in terms of the percentage of students at the highest proficiency levels, Canada ranked third, just below Finland and Korea and just ahead of Japan; Ontario fell just below the Canadian average.

These international results, given relatively little media attention in the province, fly in the face of what has been concluded from EQAO’s provincial test results. As it turns out, furthermore, the provincial tests were originally designed by EQAO with Level 2 performance as the acceptable level. Just before their first administration, the Minister of Education of the day unilaterally declared that level 3 would be the acceptable level.

This story from one province encapsulates a great many of the challenges faced in getting test-based accountability right. It is a story, for example, about arbitrary standard setting, political rather than educational motivations for standard setting, wildly improbable expectations for progress, misdiagnosis of the “problem” to be solved through greater accountability and much more.

Under circumstances partly illustrated by this story, it is not surprising that many educators are opposed to many of the features of today’s approaches to test-based accountability. For example, Ontario’s elementary teacher federation has declared outright opposition to any form of provincial testing. Several years ago, British teachers, with substantial parent support, refused to participate in the government’s testing program for young students; the British government has recently announced that it will cancel the program.

It would be surprising if most experienced teachers in Saskatchewan were not well aware of the many misuses of large-scale provincial and state testing programs, even if the province has, to date, avoided implementing some of the features giving rise to greatest controversy, for example, every pupil testing, along with school and student level reporting of results.

### 4.3 Reasons For Caution

For teachers and administrators, resistance to testing is a rational act which can be justified on both philosophical and pragmatic grounds. Philosophical reasons have mostly to do with conflicting views about the purposes for schooling (Foster, 2004) (e.g., Kohn, 2001). Arguing that advocates of accountability view such purposes as serving primarily economic interests, those who oppose test-based accountability claim that this orientation to accountability:

- is part of a technological-rational view of the world and how to make it better embedded almost entirely in economic perspectives;
- provides means of gaining control over the public agenda through the use of statistics and numbers about things as a means of influencing policy;
- has fuelled a substantial erosion in the confidence and support of the public for public schooling (although it remains amazingly robust considering the persistence and duration of this effort).

Schooling, a great many educators will argue, while certainly serving economic interests, is about much more. It is also, for example, about participating in a democratic community, discovering how to live a more satisfying life, developing a deeper understanding about the world in which one lives, and becoming better able to participate in the human conversation.

Many current approaches to test-based accountability diminish opportunities to pursue such goals by narrowing everyone's attention to the development of tool skills (it should be noted that Saskatchewan's testing program, however, unlike others in the country, also assesses critical/creative thinking).

In reality, the association of greater accountability for public schools with economic purposes is as much symbolic as real; this association arises primarily from the *sources* of advocacy for greater accountability – often the business community, conservative, pro-business governments and economic theories of human motivation - rather than changes in the goals of education promoted by such advocates. Admittedly, however, some of the more conservative advocates of test-based accountability for schools are not much interested in equity as a goal for public schooling. For example, equity was dropped from Ontario's statement of goals for the *Common Curriculum* six months after the Harris conservative government came to power. And it is certainly the case, as we shall see below, that test-based accountability policies advocated by economically - preoccupied governments have unintentionally narrowed the curriculum actually taught in schools.

These philosophical objections to test-based accountability are more or less compelling depending very much on one's value system, political leanings, or world view. But there are a significant number of additional reasons to be cautious in approaching the development (or extension) of a test-based accountability system that are much more pragmatic in nature. Nine of the most common reasons are described below.

### ***Narrowing of academic goals.***

Most large-scale testing systems in North America confine their focus to math and language with occasional forays into science. Only in relatively rare cases, Kentucky is a case in point, have efforts been made to test students in most areas of the curriculum. Technical measurement challenges, lack of resources and concerns about the amount of time for testing explain the typically narrow focus of large-scale testing programs. If what gets measured is what gets taught - and this generally seems to be the case - the press of such tests is frequently to narrow the curriculum (e.g., McNeil, 2000). Teacher and administrators regularly express concern over the loss of time to spend on social studies, the arts, music and other worthwhile areas of the curriculum (Hoffman, Czop, & Paris, 2001).

### ***Little value for diagnostic purposes.***

Many large-scale testing programs – Ontario is an example - are designed to provide reliable results for large groups of students. So results aggregated to district and provincial levels are likely to be reliable. But as the number of students diminishes, as in the case of a single school or even a small district, few testing systems claim to even know how reliable are their results. The likelihood, however, is that they are not very reliable (Wolfe, Childs, & Elgie, 2004); this means, for example, that the results might be quite different for a class of students writing the same test two weeks apart.

Lack of test reliability greatly increases the chances that areas of student performance in need of greater attention will be misdiagnosed. But in spite of this risk, it is common for provinces to recommend that school – level results be used to determine which student outcomes are in need of greater attention in the school's improvement efforts. Often this recommendation is softened by referring to the importance of also using supplementary data. But since it is the provincial data that are used to hold schools accountable, they hold much greater weight for priority setting in most schools.

“

*But there are a significant number of additional reasons to be cautious in approaching the development of a test-based accountability system*

”



Alberta's response to this problem is to provide districts with supplementary testing resources to use for diagnostic purposes. And British Columbia is now able to track individual students (their achievement profiles along with considerable other information) through their school careers, a procedure that addresses many of the problems associated with test reliability which have been raised here (Anderson, MacDonald, & Sinnemann, 2004); Ontario is in the process of developing something quite similar.

### ***Little guidance for improvement.***

More than 20 years ago, in reference to test-based accountability, the U.S. Office of Educational Research and Improvement noted that:

Accountability is a blunt tool unless policymakers, educators and the public have information that allows them to determine the likely sources of a problem and find clues about how to fix it... accountability systems should do more than simply collect test data... (quoted in Adams and Kirst, 1999, p. 477)

This was an argument for accountability systems to expand the data which they collected beyond just student performance to those features and conditions of schools likely responsible, directly or indirectly, for such performance. Armed with this more extensive information, educators would be much better able to locate and remediate the causes of underperformance on the part of students. Few test-based accountability systems collect such data but there are exceptions. For example, U.K. schools are data rich (Earl, Watson, Levin, Leithwood, & Fullan, 2003), as a result of the efforts of local education authorities, national school inspectors and fairly active research programs sponsored by the *Department of Education and Science* and the *National College of School Leadership*.

As another example, since 2001, British Columbia has been developing a system for providing schools and districts with a rich set of information (Anderson, MacDonald & Sinnemann, 2004) in five key areas:

- Foundational skills including changes over time in students reading, writing and numeracy; this information includes comparisons with other schools and provides breakdowns by demographic groups.
- Scholastic achievement in as many as 21 subject areas in grade 12; every student must pass at least one of these exams (literacy) to graduate.
- Surveys of student, parent and staff satisfaction in areas of achievement, safety and human and social development;

- Transitions from grade to grade and from secondary to post secondary education;
- A survey of human and social development is made available by the Ministry of Education to districts for their own use.

Although the information provided by this system does not systematically describe key school conditions likely to explain variation in student learning, it is considerably more useful than student achievement information alone.

### ***Common high standards.***

In many educational jurisdictions (e.g., most U.S. states, England, Ontario), the overall goal of test-based accountability systems to which schools are being held accountable is that all students will achieve the same high standards of achievement on the same goals. Note that this is a two part goal: one part is about the goals to be achieved, the second is about the levels of achievement. The problem raised by the first part of this goal is straightforward. Increasingly, the goals of the curriculum are being homogenized across districts, provinces and states to serve the requirements of large-scale testing systems; it is too complicated to acknowledge district or regional differences in educational goals, even though such differences may well make good educational sense.

The second part of this overall accountability goal is more complicated to sort out. Should we expect all students to achieve at the same high level? This is unquestionably a laudable thing to do on both democratic and egalitarian grounds. Indeed, one might be accused of racism and other nasty things by expressing any reservations about it at all. But it is much like the aim of curing all those suffering from cancer. No one objects to the goal; there is just this small problem of not actually knowing how to do it. So, ongoing debates about whether there should be different standards for students in different programs (e.g., bilingual, special education, college preparatory, or technical-vocational programs) raise, in Newman, King and Rigdon's words, "a thicket of professional and political issues that pose continuing dilemmas for districts, states and professional organizations" (1997, p. 38).

Given our lack of knowledge on the matter, the problem is that creating identical expectations for all students places the poorer ones at a distinct disadvantage (Sadker & Zittleman, 2004); setting a standard that a student has no chance of achieving is certainly as inequitable as failing to do all that we know to maximize a student's

learning. This is because students respond to increased expectations in quite different ways. Those already achieving reasonably well often possess high levels of academic self-efficacy and, for this reason, typically respond by working harder, often increasing their achievement as the standards setters had in mind.

Of course, already high performing students are not actually the students about whom the standards setters were initially concerned. Those students, the ones already struggling with the original standards, often have low levels of academic self-efficacy and are more likely to give up and disengage than to try harder in the face of yet higher standards (Kornhaber, 2004). Evidence from several different jurisdictions which have raised standards seems to support these quite different responses to higher standards. Across the U.S., for example, Heubert notes that “Students of color, students with disabilities, English language learners, and low income students are failing state graduation tests at rates as high as 60% to 90%, leaving them to face bleak futures” (2002, pp. 742-43).

### ***Temptation to “game” the system.***

The term “gaming” recently has begun to be used to capture the responses of some educators to a testing environment in which the stakes for success or failure are perceived to be quite high. For example, some teachers “game the system” by teaching to the test when test results are used to determine whether their schools will be put on probation (Mintrop, 2004). Some superintendents have been accused of gaming the system by altering district test scores in order to avoid negative sanctions by the state.

Gaming might also entail, for example, intentionally narrowing the curriculum, helping students do better on tests by providing more assistance than official test administration procedures allow, preventing students likely to do poorly on a test from writing the test, and not promoting students unlikely to perform well on grade-level tests until their chances of doing well improve.

Most of us would argue that, on the face of it, gaming practices are dishonest or unethical. But there also are reasonable arguments, some of which we have already touched on, that some forms of test-based accountability are themselves unethical, perhaps even downright destructive for students. Therefore, neither test-based accountability systems, nor those who game them, are able to claim the high moral ground. Under these circumstances, gaming practices are unlikely to go away in response to reasoned argument or appeals to ethical behavior.

The underlying problem that gaming points to is the flawed set of assumptions about how schools work (alluded to in the introduction to this section) on which test-based accountability systems rest. High stakes for teachers, associated with many testing systems, erode the well-known intrinsic sources of teachers' workplace commitments (e.g., Dannetta, 2002; Lortie, 1975), commitments which place the welfare and growth of students at the apex of teachers concerns. These much-to-be-cherished sources of commitment are replaced, or at least have to compete with, extrinsic incentives, often the incentive to avoid negative public sanctions. Mintrop's (2004) evidence suggests that the avoidance of public shaming by having your school labeled as "failing, is a enormously strong incentive for teachers.

### ***Excessive costs for minimal gains:***

A great many test-based accountability systems, unlike Saskatchewan's, assess every pupil, even though results at the school level - never mind the individual student level - have unknown levels of reliability and, as a consequence, questionable value for either diagnosing or tracking individual student progress. Such an all-student approach is extremely expensive.

Three questions are typically raised by teachers and administrators about such costs. One of these questions is about the effects on student learning of this expenditure of money. To date, evidence (mostly from the U.S.) that such large-scale, every-pupil, testing strategies improve student learning is limited in amount and has demonstrated mixed results.

Two recent studies have reported positive associations between high stakes testing and student achievement. Roderick, Jacob and Bryk (2002) found that pupil test scores increased in Chicago with the introduction of high stakes testing. Using an index developed to reflect the relative strength of accountability in all 50 U.S. states, Carnoy and Loeb (2002) found a positive association between the "strength" (meaning higher stakes) of state accountability policies and achievement in mathematics. These associations were much greater for grade 8 than grade 4 students and were weak or non-existent among high school students.

Amrein and Berliner (2002), however, found that achievement scores on non-state developed tests (i.e., ACT, NAEP and SAT) actually went down over time in more than half of the 18 states with high stakes testing policies included in their study. Kornhaber's (2004) broad review of relevant evidence paints a similarly negative picture of the effects of high stakes testing. So it seems reasonable to

conclude that the effects of high stake testing policies are still uncertain, at best.

A second frequently asked question is about the “opportunity costs” of such testing. Could the same amount of money be spent on other initiatives with greater benefit for students? How do large-scale testing effects on students compare with the effects of spending the same amounts of money on other promising initiatives such as improving the quality of instruction, substantially lowering class sizes for struggling primary students, or repairing decaying school facilities? While this is an obviously important question, it has not produced many serious attempts at an answer. One reason for this neglect, and why the question may be only superficially important, is political. The question assumes that governments could or would redirect the money spent on testing if they knew of better ways to use it. Most governments, however, are very aware of the public’s overwhelming support for testing (Livingstone, Hart, & Davie, 2003) and willingness to support its costs. It is unlikely that the public would be nearly so eager to support a government’s use of the same money for other purposes.

A third frequently asked question is whether large-scale testing could be done more cheaply? In jurisdictions which test every pupil but do not use the results for promotion (e.g., Ontario), the technical answer is “absolutely”. A perfectly good estimate of mean levels of achievement across districts and provinces can be obtained using sampling procedures of various kinds - sampling students, sampling test items or both. Sampling greatly reduces the amount of data which needs to be collected and analyzed, not to mention the amount of student and teacher time consumed by test taking and administration. Curiously, one of the most frequent defenses of every-pupil testing is not that it provides better estimates of student achievement but that the tests would not be taken seriously by students if they were based on sampling procedures. But as it stands now, even though the results of much of the pupil testing in most provinces (e.g., the grades 3, 6 and 9 tests in Ontario) are “low stakes” for students, they do seem to take them seriously, or at least few people are suggesting otherwise. In any event, the problem of ensuring adequate pupil effort on tests has much cheaper solutions than testing every pupil.

### ***Unfair reporting of results:***

Even though most of those involved should know better by now, it is common for newspapers, and even some so-called “think tanks [see the recent Fraser Institute report, (Cowley & Easton, 2004)], to rank

schools based on the unadjusted raw scores of their students' performance on annual large-scale assessments. Almost identical rankings, as we now are painfully aware, can be produced by simply ranking schools on the basis of family income or other measures of socio-economic status without spending any resources at all on student testing. Student background variables routinely explain at least 50% of the variance in student achievement across schools. Yet, unadjusted school rankings are viewed by many as evidence of school "effectiveness", presumably because they do not understand the problem.

Teachers and school administrators are correct in bristling at the unfairness of this naïve and misleading form of reporting. When such ranking procedures are used, some schools making impressive progress in improving the achievement of their low SES students are, at minimum, publicly shamed while some high ranking schools making no progress in improving the learning of their high SES students become the object of public adulation.

This crude ranking procedure is more than just unfair to schools. It greatly enhances the possibilities for mis-identifying schools in need of additional assistance and neglecting schools that have stalled in their improvement efforts. What a waste of public resources! Especially when perfectly reasonable solutions to the problem are readily available. These solutions include reporting changes over time, rather than annual achievement results and statistical procedures which partial out the effects of family background variables (see Clotfelter & Ladd, 1996, for more detailed descriptions of these solutions).

### ***Confusing evidence of change over time:***

Conceptually speaking, monitoring the extent to which a school improves the achievement of its students over time is a much better reflection of a school's effectiveness than is its annual mean achievement scores. Technically speaking, however, arriving at a defensible estimate of such change is quite challenging. Simply attributing the difference between the mean achievement scores of this year's and last year's grade 3 students on the provincial literacy test to changes in a school's effectiveness overlooks a host of other possible explanations, for example:

- Cohort differences: This year's grade 3 students may be significantly more or less advanced in their literacy capacities when they enter grade 3. Such cohort differences are quite common, as any teacher will attest;

- Test differences: While most large-scale assessment programs take pains to ensure equivalency of test difficulty from year to year, there are often subtle and not-so subtle adjustments in the tests that can amount to unanticipated but significant differences in scores;
- Differences in testing conditions: Teachers are almost always in charge of administering the tests and their class's results on last year's tests may well influence the nature of how they administer this year's test (more or less leniently) even within the guidelines offered by the province or state;
- Differences in the external environment: Perhaps the weather this winter was more severe than last winter and students ended up with six more snow days - six fewer days of instruction;
- Regression to the mean: this is a term used by statisticians to capture the highly predictable tendency for extreme scores on one test administration to change in the direction of the mean performance on a second administration. So schools scoring either very low or very high one year can be expected to score less extremely the second year, quite aside from anything else that might be different.

Linn (2003) has demonstrated that these challenges to change scores become less severe as change is traced over three or four years. It is the conclusions drawn from simply comparing this year's and last year's scores that are especially open to misinterpretation.

Unfortunately, it is these year over year comparisons (comparisons required by *No Child Left Behind*, by the way) that are most commonly made by those who report achievement results.

### ***Insufficient local capacity:***

A primary motive driving the introduction of test-based accountability systems is the perception that schools are underperforming and that introducing more stringent forms of external accountability will improve their performance. However, results of research by Newman, King and Rigdon (1997) aimed at explaining the actual consequences of such systems, point to a significant irony. It takes substantial local capacity to benefit from the effects of external, test-based accountability systems. The performance of schools without such capacity is unlikely to improve. Those with such capacity already have well-developed internal accountability systems and are likely to use external accountability initiatives to modestly extend the effects of their own initiatives.

Evidence of this sort is in short supply. So we can't be sure how well

Newman, King and Rigdon's results will stand up over time. But their study was very competently conducted and provides the best evidence available at this

### 4.4 Reasons For Action

Challenges described Section 3.2 should sound a cautionary note to Saskatchewan trustees' deliberating about future test-based accountability policies in their province. Indeed, as a trustee, you might reasonably ask, "In light of all these challenges, why is test-based accountability so widespread?" Two of the answers to this question (no doubt there are others) are easily identified. First, according to the exceptionally coherent world view or ideology from which this form of accountability arose, it is almost unthinkable that it could not work (for a description more extensive than provided in this report, see Codd, 1999). Second, at the point of first developing and implementing many of today's best known test-based accountability systems, we had much less evidence than we do now about how such systems actually work, as distinct from how they ought to work. So the challenges outlined in Section 3.2 should be thought of as the "hard lessons" learned from actually doing test-based accountability in real schools and districts. Clearly, the theory and the reality often are at significant odds with one another.

“  
*Clearly, the theory  
and the reality often  
are at significant odds  
with one another.*  
”

Notwithstanding the need for a clear-eyed appreciation of the challenges facing test-based accountability systems "on the ground", however, it is possible to imagine trustees in Saskatchewan deciding to promote the extension of the province's current testing system and/or the testing system in their districts. You might argue, for example, that no matter the level of resistance from staff, accountability, generally, and testing in particular, is not going to go away anytime soon. Saskatchewan needs to catch up with the rest of the world on this matter. Some might argue, also, that student testing is and always has been a major component of how teachers and schools do business. If it is useful for holding students accountable, why not use it to hold teachers and schools accountable?

Another fairly compelling argument for extending the province's testing program is that public education is part of the larger public service competing for scarce taxpayer dollars and it commands a very large share of the pie (in 2001, the total expenditure on elementary and secondary education in Saskatchewan was slightly more than \$1.4 billion). Other parts of the public service in most provinces are being subject to similar versions of accountability. So why should education be exempt?



Finally, and perhaps the strongest reason for increasing the testing program in the province, is that one of its main intentions, at least in its more recent iterations, is to encourage schools to pay considerably more attention to students who normally don't do well (to both "raise the bar and close the gap"). The much criticised *No Child Left Behind* legislation of the Bush government, for example, explicitly intends this as its major purpose, even though serious questions have been raised about the means chosen to accomplish that end. There can be no question that many schools have not served a fairly large proportion of the student population particularly well. Dropout rates in secondary schools speak volumes about this problem. If we could find a way to provide a high quality education for many more children it would be a huge step forward. At least test-based accountability policies requiring the disaggregation of results around different student populations forces us to acknowledge the problem.

“  
*...to encourage schools to pay considerably more attention to students who normally don't do well.*

*There can be no question that many schools have not served a fairly large proportion of the student population particularly well.*

”

### 4.5 Getting It Right

Well, let's suppose the arguments in Section 4.4 carry the day and the province decides to extend its testing program, or you convince your district to do so. What would a defensible test-based accountability system be like? There is certainly no single, simple or neat answer to this question. But the "hard lessons" do provide some traction for us. And several additional matters seem worth considering, although they have yet to surface in other test-based accountability systems to this point. A defensible system would have at least **five key properties**.

***Measure at least a sample of the full range of outcomes considered to be important in the curriculum:***

This feature is designed to prevent narrowing of the curriculum. But it comes with significant challenges including, for example, the time and money to do such extensive testing if it were done on an every pupil basis, and reliably measuring the more complex forms of intellectual, social and emotional outcomes that are included in Saskatchewan's curriculum. Undoubtedly, the distribution of testing across the grades would also be useful in order to avoid the common problem in many high stakes testing contexts of teachers fleeing the grades in which testing takes place. Kentucky's test-based accountability policy includes this feature.

***Report, on a cyclical basis, changes in students' learning of these outcomes:***

This feature aims to ensure that schools are acknowledged primarily for the value they add to students' learning - that mean annual achievement scores do not become the basis for ranking and rewarding schools.

Three challenges come with this feature. One challenge is to recognize legitimately uneven rates of growth in school effectiveness (student achievement). School improvement does not occur at a constant rate; for very good reasons, it may occur in fits and starts with plateaus along the way (Hargreaves & Fink, 2004). So expectations for annual school improvement need to be downplayed in favor of monitoring changes over relatively long periods of time - at least three or four years.

There needs to be, as well, a non arbitrary basis for determining acceptable rates of progress in school improvement. Linn (2003) argues that one such basis is the actual growth rates of high performing schools with comparable resources and student populations. The law of diminishing returns alerts us to a second challenge. As schools become more effective, it becomes harder and more complex for them to continue to improve; increasing effort is required for every additional increment of improvement. Growth targets, if they are to be included in a test-based accountability system, need to recognize both of these challenges.

Finally, changes are often calculated using data comparing often non-comparable groups of student (this year's grade three cohort may be significantly different than last). One solution to this problem has been addressed by British Columbia which now tracks individual student achievement, among other variables, throughout a student's school career.

***Estimate the proportion of learning due to the efforts of the school system:***

The challenge to be addressed in this case is well known. We described it above as part of what is unfair about test-based accountability systems that do not partial out, or otherwise control for, differences between schools in student populations; these are differences, for example, in prior learning, family background, culture and ethnicity, and language. While these factors have a huge influence on mean levels of achievement, they lie outside of what is readily controllable by schools.

This challenge can be met through a variety of well developed statistical procedures (e.g. Ladd, 1996), as well as focusing on change scores rather than mean annual achievement.

### ***Calculate the costs of learning:***

At present there are no systems (as far as I know) that include this feature. But it seems an odd feature to overlook. If schools are to be accountable, surely questions about value for money must eventually be asked.

The closest that current accountability systems come to acknowledging the importance of these questions is the implementation of performance-based budgeting systems. Such systems have appeared and then disappeared in public agencies in the U.S. without much effect. But efforts in Florida, especially in education, have been more persistent. These systems attempt to link funding to outcomes, typically increasing or decreasing funding depending on schools' success in meeting student performance targets (Herrington, 1999). Schools have proven to be especially difficult organizations in which to implement such systems.

Answering questions about value for money, at its root, requires the calculation of a ratio: the amount of learning added by the school over a fixed period of time compared with the amount of money spent to achieve such learning. Performance-based budgeting systems don't do this. They examine variation in outcomes only, leaving untouched the obvious link between inputs and outcomes.

While something like a learning-per-dollar ratio would provide a descriptive level of accountability not yet associated with test-based accountability, justification would call for some form of comparison. One would want to know whether one was getting good value for our money. This question might be addressed by borrowing a solution from current methods used in research on educational finance (e.g., Conley & Picus, 2003). These methods aim to determine the amounts of money needed to actually achieve policy-makers goals (calculations of "adequacy"). Essentially, the learning-per-dollar ratio of a school would be compared to the same ratio for comparable "high performing" schools. Different increments of learning might be justified by different levels of available resources. Or those responsible for funding schools could be held accountable for providing the same inputs enjoyed by comparable, high performing schools.

This feature of a defensible test-based accountability system is what makes it "reciprocal", a term signifying the widely distributed nature of responsibility for student performance. Ideally, for example: provincial governments would fund schools at levels adequate to

achieve the goals to which those governments aspire; trustees would enact specific policies and allocate money in support of district priorities; school and district leaders would provide managerial systems, vision, motivation, opportunities for capacity building and rewards for good performance; teachers would provide instruction and many other forms of support for students; parents would provide their children with expectations for success, adequate nourishment and a rich educational culture in the home.

***Develop local school capacity for internal accountability:***

Building on Newman, King and Rigdon's evidence, local capacity consists of explicit school-wide standards focused on student performance, mechanisms for collecting and reviewing relevant information, and a culture of peer pressure among teachers that served as potentially important consequences. When these conditions prevailed in schools, they cultivated a "*widespread consensus around professional norms and offered a focus for collaborative activity, which stimulated inquiry and a search for additional professional knowledge among staff*" (1997, p. 55). Schools with such strong internal accountability processes also looked to the external environment for signals and resources to help them define their own standards and determine the kind of information they needed to collect.

Internal school capacities of this sort, these researchers suggest, is fostered by districts and provinces when they (a) set expectations for individual schools to develop these capacities and (b) provide the staff development needed to meet those expectations.

# 5

## **SIX EXAMPLES OF WHOLE ACCOUNTABILITY SYSTEMS**

As we noted in Section 2, while there are distinctly different approaches to accountability, most educational systems mix and match approaches, typically embedding them in more comprehensive reform strategies. The outcome of this largely political process is a set of features or properties shared by many performance-oriented approaches to reform (e.g., Fuhrman & Massell, 1992; Leithwood, Jantzi, & Mascall, 2002; Newman, King & Rigdon, 1997; Smith & O'Day, 1991). These features, modestly extending those associated more narrowly with test-based accountability alone in the previous section, include:

- A centrally determined, unifying vision, and explicit goals for student performance based on the vision;
- Standards for judging the quality or degree of success of all students;
- Curriculum frameworks and related materials for use in accomplishing the goals set for students;
- Coherent, well integrated policies that reinforce these ambitious standards;
- Information about the organization's, and especially the students', performance;
- A system of finance and governance that devolves to the local school site responsibility for producing improvements in system and student performance;
- An agent that receives information on organizational performance, judges the extent to which standards have been met, and distributes rewards and sanctions, with significant consequences to the organization for its success or failure in meeting specified standards.

There are, of course, cases of accountability-driven reform efforts which do not include all seven of these elements, just as there are cases that begin less comprehensively but gradually add all or most of these feature. This section illustrates, more specifically, the nature of these seven elements by describing their appearance in the context of six whole cases of large-scale, accountability-oriented educational reform. Also described is the range of strategies that have been used to implement such reforms and the outcomes of these efforts after about a decade of implementation, in most cases. Six jurisdictions have been selected to serve these purposes – Kentucky,

California, Victoria (Australia), New Zealand, Chicago and England (specifically the initiatives associated with its National Literacy and Numeracy Strategies). These examples were chosen because of their relative maturity, the availability of data about them, and their incorporation of most of the features associated with accountability-oriented reforms.

### 5.1 Context For Accountability

While the purposes for this section do not require detailed knowledge of the contexts out of which the reforms arose, a few elements of that context are vital to understand. A government-wide sense of fiscal crisis was central to the reforms in both Victoria and New Zealand. This shaped the visions established for these reform efforts: improving the quality and equity of education had to be accomplished in the company of both increased efficiencies in the use of resources, and significant reductions in overall expenditures.

Saving money was not an objective driving reforms in either Chicago and Kentucky. These initiatives were stimulated, rather, by widespread concerns about the nature and quality of the “delivery system”: Chicago schools were believed to be drowning in a sea of stifling central bureaucracy; the Kentucky school system was declared unconstitutional by the state supreme court for reasons which included inequitable distribution of resources, not to mention nepotism and graft.

The California reform effort, while prompted by a belief that instruction could be improved in response to developments in learning theory and instructional technology, took place during a period of general dissatisfaction with public schools as reflected in the publication of *A Nation At Risk*, and subsequent reaction to the report.

England’s National Literacy and Numeracy Strategies (NLNS) were prompted by low levels of achievement on international tests and a widespread belief that England’s economic competitiveness required an education system that would raise overall achievement levels, and in particular, substantially reduce the gap between those doing well in the system and a very large proportion of children doing very poorly. Education was the main priority for Tony Blair’s labor government when it first came to power and NLNS was its main strategy for addressing this priority.

“  
...*differing levels of perceived crisis and various attributions of the causes of the crisis.*  
”

So all six examples were “revolutions” prompted by differing levels of perceived crisis and various attributions of the causes of the crisis.

### 5.2 Vision and Goals

Visions that have driven performance-based approaches to large-scale reform in the six cases are both educational and financial in nature. Educational visions vary considerably in their ambition and the extent to which they are forward looking. Chicago's vision, for example, was to increase student achievement, viewed quite traditionally, to the national average. While this is an ambitious enough goal in the context, it may not be as ambitious as, for example, the development of much higher levels of achievement for all students, as in the case of Kentucky and California. Both of these states had conceptions of achievement that encompassed both traditional and non-traditional goals. More equitable achievement on the part of the student population as a whole, was part of the educational visions of Kentucky and Victoria. New Zealand and Victoria explicitly included, as part of their educational visions, better preparing children for the 21st century. England's NLNS pointedly concerned itself with significant improvements in the "tool skills" of literacy and numeracy, aiming to both raise overall levels of achievement in these areas, as well as close the gap between academically "have and have not" students.

“*Visions that have driven performance-based approaches to large-scale reform in the six cases are both educational and financial in nature.*”

Visions unrelated to student achievement focused on either using educational resources more efficiently, as in the case of Chicago, Victoria and New Zealand, and/or reducing overall education spending, also the case in Victoria and New Zealand. Both Chicago and Victoria are examples of sites in which a more democratic form of school governance was important enough to be counted as part of their vision for the reform. Several reform efforts included a more specific set of goals associated directly with their vision. This was the case, for example, in Kentucky, where the state's General Assembly established six categories of goals toward which to drive the reform effort.

In sum, the long-term directions for the six cases of reform were in some cases focused exclusively on education, in some cases primarily concerned about reducing public resources spent on education, and in other cases a combination of both. In New Zealand and Victoria, the reform efforts took place in a context of financial crisis, helping to explain the concern for bringing under control or reducing expenditures. For the three reform efforts carried out in the U.S. context, spending less on education was not a priority. Indeed, in Kentucky, the goal was to increase the overall allocation of money to schools and to redistribute it more equitably. England actually increased its total expenditures on primary education by about five percent over the initial four year period of NLNS implementation (Earl et al., 2003).

### 5.3 Standards

As the benchmarks against which to judge performance, standards are a necessary component of performance-based approaches to reform. All six cases included the setting of performance standards for all students. Indeed, England set a series of ever more ambitious targets of student performance on its “Key Stage 2” test and declared that the minister responsible for education would resign if they were not met.

In addition, standards for teacher certification were part of the reforms in Kentucky and New Zealand. And three jurisdictions, Kentucky, Victoria, and New Zealand, also established teaching and administrator standards. California adopted standards developed by the National Council of Teachers of Mathematics. England developed a set of standards for both teachers and school administrators.

“  
*...standards are a necessary component of performance-based approaches to reform.*  
”

### 5.4 Curriculum Frameworks and Related Material

Five of the six examples of performance-based reform developed curriculum frameworks and content guidelines which further refined and specified student standards. These frameworks varied enormously in terms of their detail. Kentucky, for example, assuming that further specification would occur at the district level, specified state standards at a fairly general level. Both England and New Zealand developed a national curriculum which specified learning outcomes against which students could be measured. In New Zealand, this curriculum has become increasingly more detailed since its original inception. California curriculum frameworks also have been revised several times between 1985 and 1999; these revisions reflecting changing political and educational influences.

In most cases, curriculum frameworks are accompanied by additional guidelines intended to be of direct use by teachers. For example, England has produced substantial amounts of additional, targeted curriculum materials to assist teachers implement NLNS; New Zealand has developed support documents in languages, mathematics, science, technology, social studies, and health and physical well-being. California developed “replacement units” to supplement what were considered to be inadequacies in available textbooks.

“  
*In most cases, curriculum frameworks are accompanied by additional guidelines intended to be of direct use by teachers.*  
”



Chicago stands out among the five cases as having not devoted resources to the development of curriculum frameworks at the state level. In Chicago, some external agencies attempted to develop “teacher-proof” curricula that anyone could present to students to make up for the state’s lack of action on this matter.

### **5.5 Policies: Coherence and Integration**

Another necessary part of performance-based reform is the creation of policies designed to govern the reform process which are coherent and integrated. Typically this is what is meant when such reforms are referred to as “systemic”. While each of the six cases demonstrates some attention to coherence across policies, this concern is much more comprehensive in some cases than in others. England is arguably the most comprehensive, as well as ambitious. By now policies for resource allocation, governance, personnel assessment, both teacher and leader development and a host of less obvious policies have been carefully aligned in support of NLNS implementation. California has demonstrated a strong concern about the relationship between curriculum policies, assessment policies, and policies governing textbook adoption. Much less attention, however, was devoted to policies framing the use of curriculum frameworks and textbooks. Kentucky, in contrast, illustrates a very comprehensive approach to policy coherence. Not only was a significant effort devoted to ensuring the coherence of curriculum and assessment policies, for example, but the Kentucky Education Reform Act also provided the state with policies governing finance, governance, incentive systems and capacity-building strategies aimed at supporting the implementation of the curriculum and assessment policies.

Both New Zealand and Victoria also illustrate highly comprehensive approaches to policy integration and coherence. In Victoria, for example, the starting point for this coherence is four carefully integrated policy frameworks, concerned with curriculum, people, resources, and accountability. These four frameworks are intended to work together toward the vision for education held by the state.

The six examples as a whole, then, illustrate the range from weak to very strong efforts to align policies around implementing the vision for education.

## **5.6 Performance Information**

The six cases also vary considerably in terms of how fully-developed was the performance information system included as part of the reform effort. Kentucky, Chicago, and California limited their collection of performance information to student achievements of various sorts. Even across these three states, the extent of performance information development varies widely. In the case of Chicago, performance information from students was collected through annual tests of elementary and secondary students in both reading and mathematics. Kentucky, in contrast, collected data from students in grades 4, 5, 7, 8, 11, and 12 in seven different subject areas.

California's approach to collecting performance information evolved considerably over a decade. Initially such data were collected through the California Assessment Program, using a matrix sampling technique. This program was canceled in 1990, and replaced by the California Learning Assessments System. This was an ambitious attempt at performance-based assessment, but eventually had to be canceled in the face of both technical difficulties and political opposition to performance-based approaches to assessment. Indeed, the California experience with performance-based assessment mirrors that of Kentucky. Initially, Kentucky's tests were performance-based, eventually evolving to mixed-question types and, finally, to complete elimination of performance events for accountability purposes. This, too, was in response to criticisms regarding the validity of such tests.

While the U.S.-based cases largely restrict their performance data to student achievement, both Victoria and New Zealand collect a broader array of data. These include, in addition to achievement data, data about school performance against goals established for the organization in its charter. In Victoria, a cooperative research project was established to monitor the implementation of the Schools for the Future project over five years. New Zealand annually collects baseline achievement data on a representative, 3% sample of years 4 and 8 students. An annual national report on school performance is also produced by the Educational Review Office and individual schools publish annual accountability reviews.

England's student performance information system depends heavily on its Key Stage tests. But the English education system is data rich with respect to other types of performance, especially the performance of school organizations. Such information is provided by a national inspectorate, which conducts periodic site visits to schools armed with quite specific criteria on which its observations, evaluations and recommendations are based.

### 5.7 Finance And Governance

Kentucky and Chicago are noteworthy in their efforts to redistribute funding away from the wealthiest schools and districts in order to achieve more equitable funding for all students. This redistribution of money, however, is still controlled by the state. Both England, Victoria and New Zealand have encouraged schools to move toward control over a very large proportion of their total funding. New Zealand, for example, has encouraged schools to adopt the “fully funded option” that includes control over teachers’ salaries, an option so far selected by about a third of the schools in the country. That option is also the long term goal in Victoria with the “school global budget” an interim step. According to this policy, schools are allocated a base grant calculated on a per capita basis, with additional funds depending on the profile of students in the school.

Victoria and New Zealand, approached the funding of schools in the face of a serious crisis in government debt. As a consequence, in Victoria for example, the state education budget was cut by 15% over five years.

In each of the five reform efforts there has been careful targeting of funds directly to reform initiatives. For example, California gave districts five dollars per student for administering the California Assessment Program, and thirty dollars per student for implementing the school improvement program related to the implementation of curriculum frameworks. Funding in Kentucky was provided for all programs included as part of the Kentucky education reform. As we noted above, England’s expenditures on primary education increased about five percent over the first four years of NLNS implementation and much of this was targeted to services for struggling schools.

On the matter of governance, in the majority of the five cases significant changes were made in governance structures at both the central and local levels. These changes often were driven by a widely-adopted governance model which includes central control over policy development and monitoring, and greater local control over policy implementation and operations. This model has been well enunciated in the context of the English, New Zealand and Victoria cases, which have more fully devolved responsibility to local schools than is the case in the three American examples. In New Zealand, for example, schools are relatively independent entities with their own charters, governed by their own local councils, and responsible to their local

“  
...widely-adopted  
governance model  
which includes central  
control over policy  
development and  
monitoring, and greater  
local control over  
policy implementation  
and operations.”

communities. Central oversight is restricted to various review initiatives undertaken from time to time by the central government.

The move toward greater devolution of operational authority to the local school level was stimulated, in the cases of especially New Zealand, Victoria and Chicago, by considerable criticism of the deadening effect of excessive bureaucracies in those systems. These central structures have been modified considerably in each of these three cases to be more directly responsive to school needs and to eliminate many of the control functions which they carried out in the past. In Victoria, for example, out-of-school bureaucracy was reduced by 60% with staff in the central bureaucracy reduced from 2300 to 600. New Zealand reduced its governance structure from three to two levels, eliminating the entire middle level of regional boards (LEAs)

### 5.8 Agent Receiving Information, Distributing Rewards And Sanctions

In Victoria, Kentucky and New Zealand, special agencies were established with specific responsibilities for accumulating performance information, but the central governments maintained responsibility for distributing rewards and sanctions in these cases. The mayor's office in Chicago took control of the Chicago District School Board, initially. In the 1995 revision of this reform effort in Chicago, the Mayor's office also had responsibility for sanctions against non-performing schools and the provision of extra resources to improve schools not up to standard. In Chicago, New Zealand and Victoria, principals were given greater authority to sanction underperforming teachers. Local school councils were awarded power to hire and evaluate principals in Chicago and New Zealand. Local governing bodies are responsible for school development in England. But the government's own Standards and Effectiveness Unit monitors the performance of schools across the country and plays the lead role in determining rewards and sanctions for performance.

“  
...special agencies  
were established with  
specific responsibilities  
for accumulating  
performance  
information  
”

In all of these efforts to redistribute authority, district authority has been weakened, and in the case of New Zealand eliminated entirely.

# 6

## TOWARD A MODEL OF PROFESSIONAL ACCOUNTABILITY IN SASKATCHEWAN

### 6.1 Introduction

This section of the report recommends adoption of a professional approach to accountability in Saskatchewan and provides further details about what such an approach might entail. Such an approach is one of four quite different alternatives described in some detail in Section 3 of this report: market, decentralization, professional and managerial approaches. Rarely, however, have market, decentralization or management approaches lived up to their expectations, as evidence reviewed in Section 3 and elsewhere in the report indicates

“  
*...adoption of a professional approach to accountability*  
”

While only modest amounts of evidence are available about the effects of professional approaches to accountability, that evidence is generally positive. Furthermore, the assumptions and broad features of a professional approach to accountability are quite consistent with what we know from other sources about how to improve student learning.

Whereas the other three approaches to accountability are premised on the need for greater motivation on the part of educators, a professional approach rests on the belief that teachers and administrators are already highly committed to the welfare and learning of their students. Such commitment is not in doubt among educators in this province. Professional approaches to accountability, then, build on this foundation of commitment by investing heavily in the initial and continuing development of educators' expertise rather than treating such expertise as an inevitable outgrowth of greater extrinsic motivation.

“  
*...Professional approaches to accountability, then, build on this foundation of commitment by investing heavily in the initial and continuing development of educators' expertise.*  
”

This focus on enhancing the expertise of provincial educators should not be interpreted as a criticism of their current capacities. Rather, it reflects the value of continuous learning widely shared in a professional learning community, a value fuelled by the belief that the future world in which our children will live and work is certain to be a different and a much more complex world than it is presently. Schools will need to adapt their practices accordingly and successful adaptation will inevitably entail the acquisition of new capacities on the part of those who work in schools.

## **6.2 Basic Features of a Professional Approach To Accountability**

Section 3.4 briefly described the main features of a professional approach to accountability drawing on the work of Wise and Liebbrand (2003) undertaken for the Gates Foundation in the United States. We continue to be influenced by their work in this section, although what is proposed for Saskatchewan is certainly different in detail than what Wise and Liebbrand described for the United States.

Figures 1, 2, and 3 summarize the main features of a professional approach to accountability for the province. Figure 1 specifies accountabilities at the school level while Figures 2 and 3 outline accountabilities for school divisions and the province respectively. As in the case of Wise and Liebbrand, these figures answer four questions (Who is accountable? For what? To whom? With what measures?) about all of the main stakeholders in the education system.

Figures 1, 2 and 3 illustrate the nature of a *reciprocal* accountability system in very general terms. Should the province move toward the adoption of such an approach, decisions embedded in Figures 1 to 3 would need to be carefully revisited and developed in considerably greater detail. The approach to reciprocal accountability described in the Figures, then, is a starting point for serious discussion not a finished product.

Figure 1 describes the potential features of a reciprocal, professional approach to accountability at the school level. The accountabilities of parents, students, teachers and administrators are all described: what they are accountable for, to whom and how those accountabilities might be assessed. Figure 1 suggests, for example, that although teachers are accountable for implementing professional teaching standards, administrators are accountable for providing the resources and other forms of assistance teachers need for that job. Furthermore, while students are accountable for achieving standards outlined in the curriculum, parents are accountable for providing an environment which encourages, and as far as possible, enables success at school.

Figure 2 steps back from the school and outlines the potential nature of accountabilities at the level of the school division, accountabilities of both the school board and the senior administration. By way of example, one of the accountabilities of senior administration spelled out in this figure is the alignment of school division policies and procedures (e.g. funding policies, curriculum policies, assessment policies, performance appraisal policies) in

support of programs and other initiatives to help students achieve curriculum standards. Such alignment is a major contribution to the effective work of school administrators and teachers, ensuring that policies and procedures actually support and make possible, rather than standing in the way of, initiatives in schools aimed at achieving standards specified for students.

Finally, Figure 3 outlines accountabilities for provincial roles and agencies – the Minister of Education, Saskatchewan Learning, Faculties of Education, Professional Associations and a College of Teachers. A major accountability at this level, by way of illustration, falls to the College of Teachers. While teachers are held accountable for implementing professional teaching standards, administrators for helping them acquire the capacities those standards describe, and district staff for ensuring a congenial policy environment, it is up to the College to develop these standards in the first place. Likely, it is also up to the College to ensure the use of such standards in the initial training, licensure and continuing development of teachers in the province, as well.

Seven features embedded in Figures 1, 2 and 3 are central to professional approaches to accountability:

### **Student Standards.**

The ultimate benchmark for judging progress in a professional accountability system is the extent to which students make progress in achieving whatever goals, objectives or standards are included in the provincial curriculum. Adopting a professional approach to accountability in the province does not automatically trigger the need for provincial curriculum revision. But the province does need to satisfy itself that the expectations for student growth through formal education are clearly outlined and sufficiently well-specified to serve as a source of practical guidance to educators and to allow for assessment of their achievement.

### **Assessments Of Student Progress.**

Section 4 of this report offered a fairly extensive exploration of the challenges associated with assessing student progress. While these challenges are not rehearsed here, we do note that they remain to be worked on within a professional approach to accountability. That said, a professional approach does not rely on them as exclusively as some other approaches to accountability. So

“  
...other important  
evidence should be  
collected about the  
extent to which  
standards of practice on  
the part of teachers and  
administrators, other  
obligations are being  
carried out  
”

failure to address the challenges in a fully satisfactory way (not very likely in the near future) will not be as critical as it would be within a different approach accountability. For example, as Figures 1 to 3 indicate, other important evidence should be collected about, for example, the extent to which standards of practice on the part of teachers and administrators, or other obligations (parents providing expectations supportive of success at school) are being carried out by those who are held accountable for them.

***A College Of Teachers.***

Several provinces have established professional colleges of teachers at least partly reflecting the licensing and monitoring agencies central to other well established professions. In the case of Ontario's college, for example, the degree of control by teachers over the college, however, is quite limited in comparison with the comparable agencies in medicine or law. Furthermore, Wise and Leibbrand claim that "... 43 states in the U.S. have created state standards or practice boards, with the majority of the members being practicing teachers" (2003, p. 32)

A professional model of accountability would be incomplete without such a board or college to govern the affairs of the profession and such a board might well replace the several boards now part of Saskatchewan Learning which currently oversee matters concerning the training and accreditation of teachers. One of the College's first tasks would be the establishment of professional standards for its members.

“  
*A professional model of accountability would be incomplete without such a board or college to govern the affairs of the profession*  
”

***The development and use of professional standards.***

In Figures 1 to 3, these are standards for teachers, and administrators. But standards could also be specified for such other roles as trustees and parents. If Saskatchewan were to adopt a professional approach to accountability along the lines described in this section, adopting, adapting or developing such standards would be a critical initial task to carry out.

A considerable amount of work has been carried out to develop such standards in other jurisdictions by this point. The products of much of this work have been alluded to throughout this paper. So building on this work rather than inventing an entirely new set of standards would be seem to make good sense.



***Policies, procedures and experiences “wrapped around” each set of standards aimed at helping build the capacities indicated by the standards.***

Once professional standards have been developed, the task becomes one of orienting professional learning experiences toward their development in both new and experienced teachers and administrators; this task is entirely consistent with the norms and beliefs of a professional learning community. In addition to the “natural” or on-the-job opportunities available for learning within a professional learning community, the province already has formal structures designed for professional learning that could be more directly focused on the standards. Whether or not these existing structures are sufficient is a question that needs to be asked. The initial building of professional capacities is the work of faculties of education and at least a significant portion of what Saskatchewan’s faculties of education do could be aligned with a set of provincial teaching standards. Nonetheless, after Wise and Liebbrand, a sequential conception of teacher development and licensing might be needed to significantly increase professional capacities across the province. Such a sequential conception might include, for example:

- pre-service education keyed to the province’s teaching standards as a minimum;
- followed by a 10 month intern period; and
- an performance requirement to be met in order to acquire a teaching license.

Responsibility could then transferred to schools and professional associations for continuing professional development as is presently the case. But consideration likely should be given to extending the province’s resources in this area. For example, England has created the *National College of School Leadership* in recognition of the limited progress being made in achieving their leadership standards by local education authorities and private service providers alone. Similarly, in the U.S. many training centres for both teachers and administrators, some state sponsored and some associated with universities or unions, have sprung up for the same reasons.

“  
*...the task becomes one of orienting professional learning experiences toward their development in both new and experienced teachers and administrators;*  
”

***Methods for assessing the extent to which such capacities have been acquired and are being applied.***

Figures 1-3 identify both school and school division performance appraisal systems as the primary means of measuring the extent to which administrators and teachers have acquired the capacities and are practicing in the manner prescribed by the professional standards. This is a logically sensible function for such performance appraisal systems and it would be hard to justify their continuing existence in a professional accountability context were they not able to adequately fulfill this function. But performance appraisal systems are not generally noted for contributing much to professional growth, so a careful review of such systems in the province is likely warranted and some significant “upgrading” of existing systems may be required. For example, the leadership standards developed by the U.S. Chief State School Officers, used in many U.S. states (see Section 3.4) are now accompanied by a set of specially designed assessment instruments developed by Educational Testing Services. These instruments offer considerably more precise information about the possession of leadership standards than would the typical performance appraisal system.

***Reciprocal accountability relationships among all stakeholders.***

While it is common practice for those in schools to be the object of many current accountability policies, Figures 1, 2 and 3 begin to specify accountabilities for all stakeholders. As we noted above, in reference to Figure 1, teachers are held accountable for acquiring the capacities specified in a set of teaching standards. But a College of Teachers is accountable for identifying those standards to begin with and school administrators are accountable for assisting teachers acquire the capacities outlined in the standards.

The rationale for reciprocity in an accountability system can be found in our earlier discussion of the importance of not holding people accountable for things over which they have either no, or only partial, control. As we noted in Section 2.4, this is not only unethical, it defeats the purposes that are served by coordinated action. All stakeholders have a contribution to make, directly or indirectly, to the quality of a student’s education. So all stakeholders should be held accountable for providing whatever it is of value they are both obliged and in a position to add.

“  
*All stakeholders have a contribution to make, directly or indirectly, to the quality of a student’s education.*  
”

The importance of reciprocity is easily demonstrated. As we noted earlier in reference to Figure 1, while teachers can be held fully accountable for the nature of instruction used in their classrooms, much of what students learn from such instruction depends on the attitudes and expectations of students concerning the value of schooling. Parents have a large part to play in the formation of such attitudes and expectations. Parent accountability for the development of positive attitudes and expectations might well be part of a reciprocal accountability system.

# 7

## CONCLUSION

Widespread government efforts to hold schools more publicly accountable are almost always well intentioned. Parents are right to be concerned, for example, about how well schools are preparing their children for life in increasingly complex and pluralistic social environments. Communities understandably worry about whether enough is being done by schools to instill in their youth the sense of shared regard and civic responsibility demanded of a functioning democracy. And private sector employers cannot be faulted for wanting schools to be more attentive to the preparation of future employees with the insights, problem solving skills and other capacities needed for their success in the Darwinian context of a world-wide marketplace. We should not be surprised that parents, communities and employers demand vigorous efforts by their elected representatives to address these concerns.

While schools have proven extraordinarily adept at making adaptive, incremental changes within the boundaries of a widely shared understanding of what schools are and do, they have never been expected to change as rapidly and fundamentally as they are being asked to do at this time in their history. Government efforts, in many parts of the world, to hold schools more publicly accountable are part of a larger reform effort aimed at such fundamental change. It is clear from the evidence reviewed in this paper, however, that such efforts to hold schools more publicly accountable typically have quite uneven consequences for students and professional educators.

The contemporary accountability movement in education, which began some 10 to 15 years ago, was guided by a very comprehensive political ideology. But there was almost no empirical evidence, at that time, about the actual effects on students, schools and professional educators of policies and practices recommended by that ideology. We now have a considerable amount of such evidence and paramount among the lessons which it teaches us is that:

- many of the most ideologically popular accountability policies actually exacerbate the problems they were designed to solve;
- almost all approaches to holding schools more publicly accountable require considerable local capacity and engagement if they are to accomplish their desired goals;

- approaches to accountability focused on building the capacity of those in schools and districts to work smarter (professional approaches) have much more potential to be effective than approaches focused only on motivating those in schools and districts to work harder; and, finally,
- in order to bring the full set of resources needed by the education system to approximate the ambitious goals of large-scale reform, a systems approach to accountability is required, an approach which specifies the “reciprocal” accountabilities of all those who have something to contribute to the solution.

Those deliberating about the most suitable forms of educational accountability for Saskatchewan would do well to build on these lessons.

**Figure 1**

***Possible Features Of A Professional Approach To Accountability In Saskatchewan:  
The School Level***

<b>Who Will Be Accountable?</b>	<b>For What?</b>	<b>To Whom?</b>	<b>With What Measure?</b>
School Administration	<ul style="list-style-type: none"> <li>• Acquiring and implementing capacities identified in leadership and administrative standards</li> <li>• Helping staff develop and implement capacities identified in professional teaching standards</li> </ul>	<ul style="list-style-type: none"> <li>• Director</li> <li>• Teachers</li>   <li>• Director</li> <li>• Teachers</li> <li>• Parents</li> <li>• Students</li> </ul>	<ul style="list-style-type: none"> <li>• Performance appraisal system</li>   <li>• Performance appraisal system</li> <li>• Local and provincial student assessment data</li> </ul>
Teacher	<ul style="list-style-type: none"> <li>• Acquiring and implementing capacities identified in professional teaching standards</li> </ul>	<ul style="list-style-type: none"> <li>• Principal</li> <li>• Students</li> <li>• Parents</li> </ul>	<ul style="list-style-type: none"> <li>• Performance appraisal evidence</li> <li>• Local and provincial student assessment data</li> </ul>
Student	<ul style="list-style-type: none"> <li>• Meeting student standards</li> <li>• Assisting peers achieve student standards</li> </ul>	<ul style="list-style-type: none"> <li>• Parents</li> <li>• Teachers</li>   <li>• Other students</li> </ul>	<ul style="list-style-type: none"> <li>• Local and provincial student assessment data</li> <li>• Teacher and student opinion</li> </ul>
Parent	<ul style="list-style-type: none"> <li>• Explicit support for the value of student standards</li> <li>• Expectations/encouragement of child to achieve standards</li> <li>• Information to school relevant to child's learning</li> </ul>	<ul style="list-style-type: none"> <li>• School</li>   <li>• School</li>   <li>• School</li> </ul>	<ul style="list-style-type: none"> <li>• Staff/student opinion</li>   <li>• Staff/student opinion</li>   <li>• Staff/student opinion</li> </ul>

**Figure 2**

***Possible Features Of A Professional Approach To Accountability In Saskatchewan:  
The School Division Level***

<b>Who Will Be Accountable?</b>	<b>For What?</b>	<b>To Whom?</b>	<b>With What Measure?</b>
Board of Education	<ul style="list-style-type: none"> <li>• Provide resources for ongoing professional development</li> <li>• Hire Director who meets standards</li> <li>• Establish, implement and refine effective performance appraisal system for Director of Education</li> <li>• Complying with current legislation</li> </ul>	<ul style="list-style-type: none"> <li>• Senior administration</li> <li>• Community and Saskatchewan Learning</li> <li>• Community and Saskatchewan Learning</li> <li>• Saskatchewan Learning</li> </ul>	<ul style="list-style-type: none"> <li>• Existence of ongoing PD</li> <li>• Performance appraisal data</li> <li>• Existence of PAS</li> <li>• Expert judgements</li> <li>• Periodic evaluation of PAS</li> <li>• Government audits</li> </ul>
Director of Education and Central Office Staff	<ul style="list-style-type: none"> <li>• Hire district and school staff who meet or are able to acquire standards</li> <li>• Provide ongoing high quality professional development</li> <li>• Establish, implement and refine effective performance appraisal systems for teachers, school and district administrators</li> <li>• Align district policies in support of programs to accomplish student standards</li> <li>• Develop widely shared sense of direction for district</li> <li>• Develop culture of collaboration in the interests of student learning</li> </ul>	<ul style="list-style-type: none"> <li>• Board of Education</li> <li>• Board of Education</li> <li>• Central office staff</li> <li>• Board of Education</li> <li>• Central office staff</li> <li>• Saskatchewan Learning</li> <li>• Board of Education</li> <li>• Central office staff</li> <li>• Board of Education</li> <li>• Central office staff</li> <li>• Board of Education</li> <li>• Central office staff</li> </ul>	<ul style="list-style-type: none"> <li>• Staff credentials</li> <li>• Performance appraisal information</li> <li>• Existence of PD</li> <li>• Evaluations of PD</li> <li>• Existence of PAS</li> <li>• Periodic evaluations of PAS</li> <li>• Opinions of staff</li> <li>• Opinions of staff</li> <li>• Opinions of staff</li> </ul>

**Figure 3**

***Possible Features Of A Professional Approach To Accountability In Saskatchewan:  
The Provincial Level***

<b>Who Will Be Accountable?</b>	<b>For What?</b>	<b>To Whom?</b>	<b>With What Measure?</b>
Minister of Education	<ul style="list-style-type: none"> <li>• Building public and professional support for goals of public education</li> <li>• Ensuring adequate funding</li> <li>• Establishing policies capable of achieving goals of education</li> </ul>	<ul style="list-style-type: none"> <li>• The public</li> <li>• Education professionals</li> <li>• The public</li> <li>• The public</li> <li>• Education professionals</li> </ul>	<ul style="list-style-type: none"> <li>• Public opinion polls</li> <li>• Voter support</li> <li>• Educator support (surveys)</li> <li>• Level of funding relative to need and to other jurisdictions</li> <li>• Public opinion</li> <li>• Educators' opinion</li> </ul>
Saskatchewan Learning	<ul style="list-style-type: none"> <li>• Communication policies</li> <li>• Developing procedures and regulation</li> <li>• Providing basic support for district policy implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Minister</li> <li>• Districts</li> </ul>	<ul style="list-style-type: none"> <li>• Surveys of policy clarity</li> <li>• Clarity about implementing policy</li> <li>• Division staffs' opinion about feasibility of policy implementation</li> </ul>
Faculties of Education	<ul style="list-style-type: none"> <li>• Consult/advise College of Teachers on standards and assessment</li> <li>• Initial preparation keyed to standards</li> <li>• Collaborate with districts and schools on internships</li> </ul>	<ul style="list-style-type: none"> <li>• College of Teachers</li> <li>• College of Teachers</li> <li>• School divisions and schools</li> </ul>	<ul style="list-style-type: none"> <li>• Opinion of College of Teachers</li> <li>• Opinion of College of Teachers</li> <li>• Opinion of College of Teachers</li> </ul>
Professional Associations	<ul style="list-style-type: none"> <li>• Consult/advise College of Teachers on standards and assessment</li> <li>• Assist members in meeting standards</li> <li>• Help build capacity to implement policy</li> </ul>	<ul style="list-style-type: none"> <li>• College of Teachers</li> <li>• Members</li> <li>• Saskatchewan Learning</li> </ul>	<ul style="list-style-type: none"> <li>• Judgements about quality of standards</li> <li>• Opinions about value of assistance</li> <li>• Opinions about value of assistance</li> </ul>
College of Teachers	<ul style="list-style-type: none"> <li>• Creating professional standards</li> <li>• Licensing teachers who meet professional standards</li> <li>• Assess teachers' initial acquisition of professional standards</li> </ul>	<ul style="list-style-type: none"> <li>• Saskatchewan Learning</li> <li>• Saskatchewan Learning</li> <li>• Saskatchewan Learning</li> </ul>	<ul style="list-style-type: none"> <li>• Judgements about quality of standards</li> <li>• Opinion about licensing process</li> <li>• Opinion about assessment procedure</li> </ul>



# 8

## REFERENCES

- Adams, J. E., & Kirst, M. (1999). New demands and concepts for educational accountability: Striving for results in an era of excellence. In J. Murphy & K. Louis (Eds.), *Handbook of research on educational administration. 2nd edition*. San Francisco: Jossey-Bass.
- Amrein, A., & Berliner, D. (2002). High stakes testing, uncertainty, and student learning. *Education Policy Archives*(March).
- Anderson, B., MacDonald, D., & Sinnemann, C. (2004). Can measurement of results help improve the performance of schools? *Phi Delta Kappan*, 85(10), 735-740.
- Bauch, P., & Goldring, E. (1995). Parent involvement and school responsiveness: Facilitating the home-school connection in schools of choice. *Educational Evaluation and Policy Analysis*, 17(1), 1-21.
- Bay, J. R. B., & Reys, R. (1999). The top ten elements that must be in place to implement standards-based mathematics curricula. *Phi Delta Kappan* (March), 503-506.
- Beach, R., & Lindahl, R. (2004). A critical review of strategic planning: Panacea for public education? *Journal of School Leadership and Management*, 14(March), 211-234.
- Beck, L., & Murphy, J. (1998). Site-based management and school success: Untangling the variables. *School Effectiveness and School Improvement*, 9(4), 349-357.
- Bullock, A., & Thomas, H. (1997). *Schools at the centre? A study of decentralization*. London: Routledge.
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305-331.
- Chubb, J., & Moe, T. (1990). *Politics, markets, and America's schools*. Washington, DC: The Brookings Institute.
- Clotfelter, C. T., & Ladd, H. (1996). Recognizing and rewarding success in public schools. In H. Ladd (Ed.), *Holding schools accountable* (pp. 23-64). Washington, DC: The Brookings Institute.

- Clune, W. H., & Witte, P. (1988). *School-based management: Institutional variation, implementation, and issues for further research*. New Brunswick, NJ: Eagleton Institute of Politics, Center for Policy Research in Education.
- Codd, J. (1999). Educational reform, accountability and the culture of distrust. In M. Thrupp (Ed.), *A decade of reform in New Zealand Education: Where to now?* (pp. 45-53). Waikato, NZ: University of Waikato, School of Education.
- Coleman, J. S. (1986). *Equality of educational opportunity*. Washington, DC: Government Printing Office.
- Conley, D. T., & Picus, L. O. (2003). Oregon's quality education model: Linking adequacy and outcomes. *Educational Policy*, 17(5), 586-612.
- Cookson, P. W. (1994). *School choice: The struggle for the soul of American education*. New Haven, CT: Yale University Press.
- Cowley, P., & Easton, S. (2004). *Report card on Ontario's elementary schools, 2004 edition*. Vancouver, BC: The Fraser Institute.
- Cranston, N. (2000). The impact of school-based management on the primary school principal: An Australian perspective. *Journal of School Leadership and Management*(10), 214-232.
- Dannetta, V. (2002). What factors influence a teacher's commitment to student learning? *Leadership and Policy in Schools*, 1(2), 144-171.
- Dareesh, J. (1998). Professional development for school leadership: The impact of US educational reform. *International Journal of Educational Research*, 29(4), 323-333.
- Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Researcher*, 31(9), 13-25.
- Earl, L., Watson, N., Levin, B., Leithwood, K., & Fullan, M. (2003). *Watching and Learning 3: Final report of the external evaluation of England's National Literacy and Numeracy Strategy*. Toronto: OISE/University of Toronto.
- Elmore, R. (1990). Choice as an instrument of public policy: Evidence from education and health care. In W.H. Clune & J. Witte (Eds.), *Choice and control in American education* (Vol. 1: The theory of choice and control in education). New York: Falmer Press.

- Foster, W. (2004). The decline of the local: A challenge to educational leadership. *Educational Administration Quarterly*, 40(2), 176-191.
- Fuhrman, S., & Massell, D. (1992). *Issues and strategies in systemic reform*. New Brunswick, NJ: Consortium for Policy Research in Education.
- Giles, C. (1997). *School development planning: A practical guide to the strategic management process*. Plymouth, UK: Northcote House Publishers.
- Grace, G. (2002). *Catholic schools: Mission, markets and morality*. London: Routledge Falmer.
- Hargreaves, A., & Fink, D. (2004). The seven principles of sustainable leadership. *Educational Leadership*, 61(7), 8-13.
- Harrison, C., Killion, J., & Mitchell, J. (1989). Site-based management: The realities of implementation. *Educational Leadership*, 46(8), 55-58.
- Hausman, C. S. (2000). Principal role in magnet schools: Transformed or entrenched? *Journal of Educational Administration*, 38(1), 25-46.
- Henig, J. R. (1994). *Rethinking school choice: Limits of the market metaphor*. Princeton, NJ: Princeton University Press.
- Herrington, C. (1999). Performance-based budgeting for public schools in Florida. In M. Goertz & A. Odden (Eds.), *School-based Finance* (pp. 188-214). Thousand Oaks, CA: Corwin Press.
- Hess, G. A., Jr. (1991). *School restructuring Chicago style*. Newbury Park, CA: Corwin Press.
- Heubert, J. (2002). First, do no harm. *Educational Leadership*(December), 26-30.
- Hoffman, J., Czop, L., & Paris, S. (2001). High stakes testing in reading: Today in Texas, Tomorrow? *Reading Teacher*(54), 482-492.
- Kohn, A. (2001). Fight the tests: A practical guide to rescuing our schools. *Phi Delta Kappan*, 85(5), 348-357.
- Kornhaber, M. L. (2004). Appropriate and inappropriate forms of testing, assessment and accountability. *Educational Policy*, 18(1), 45-70.
- Ladd, H. (1996). *Holding schools accountable*. Washington, DC: The Brookings Institute.

- Lauder, H., & Hughes, D. (1999). *Trading in futures: Why markets in education don't work*. Buckingham, UK: Open University Press.
- Lee, V. (1993). Educational choice: The stratifying effects of selecting schools and courses. *Educational Policy*, 7(2), 125-148.
- Leithwood, K. (2001). School leadership in the context of accountability policies. *International Journal of Leadership in Education*, 4(3), 217-236.
- Leithwood, K., & Earl, L. (2000). Educational accountability effects: An international perspective. *Peabody Journal of Education*, 75(4), 1-18.
- Leithwood, K., Jantzi, D., & Mascall, B. (2002). A framework for research on large-scale reform. *Journal of Educational Change*(3), 7-33.
- Leithwood, K., Jantzi, D., & McElheron-Hopkins, C. (in press). Parent roles in school improvement planning. *School Effectiveness and School Improvement*.
- Leithwood, K., Jantzi, D., & Steinbach, R. (2002). Leadership practices for accountable schools. In K. Leithwood & P. Hallinger (Eds.), *Second international handbook of educational leadership and administration* (pp. 849-880). Dordrecht: Kluwer Academic Publishers.
- Leithwood, K., & Menzies, T. (1998). Forms and effects of school-based management: A review. *Educational Policy*, 12(3), 325-346.
- Linn, R. (2003). Accountability: responsibility and reasonable expectations. *Educational Researcher*, 32(7), 3-13.
- Livingstone, D., Hart, D., & Davie, L. E. (2003). *Public attitudes towards education in Ontario 2002: The 14th OISE/UT survey*. Toronto: Orbit Monograph, OISE/UT.
- Lortie, D. C. (1975). *School teacher: A sociological study*. Chicago, IL: University of Chicago.
- Mintrop, H. (2004). *Schools on probation: How accountability works (and doesn't work)*. New York: Teachers College Press.
- Murphy, J., & Beck, L. (1995). *School-based management as school reform*. Thousand Oaks, CA: Corwin Press.
- Newman, F. M., King, M., & Rigdon, M. (1997). Accountability and school performance: Implications for restructuring schools. *Harvard Educational Review*, 67(1), 41-74.

- Ornstein, A. C. (1983). Administrative decentralization and community policy: Review and outlook. *Urban Review*, 15(1), 3-10.
- Peters, M. (1992). Performance indicators in New Zealand higher education: Accountability or control? *Journal of Education Policy*, 7(3), 267-283.
- Prestine, N. (1999). Enabler or restrainer? Factors that determine administrator responses to systematic change initiatives. Unpublished manuscript.
- Raywid, M. (1992). Choice orientations, discussions, and prospects. *Educational Policy*, 6(2), 105-122.
- Roderick, M., Jacob, B., & Bryk, A. (2002). The impact of high stakes testing in Chicago on student achievement in promotional gate grades. *Educational Evaluation and Policy Analysis*, 24(4), 333-359.
- Sadker, D., & Zittleman, K. (2004). Test anxiety: Are students failing tests - or are tests failing students? *Phi Delta Kappan*, 85(10), 740-745.
- Scriven, M. (1994). Duties of the teacher. *Journal of Personnel Evaluation in Education*(8), 1561-1184.
- Smith, M. S., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing*. Bristol, PA: Falmer Press.
- Southworth, G. (1998). *Leading improving primary schools*. London: Falmer Press.
- Tanner, K. C., & Stone, C. D. (1998). School improvement policy: Have administrative functions of principals changed in schools where site-based management is practiced? *Education Policy Analysis Archives*, 6(6), 1-14.
- Wagner, R. (1989). *Accountability in education: A philosophical inquiry*. London: Routledge.
- Wildy, H., & Loudon, W. (2000). School restructuring and the dilemma of principals' work. *Educational Management and Administration*, 28(2), 173-184.
- Williams, R., Harold, B., Robertson, J., & Southworth, G. (1997). Sweeping decentralization of educational decision-making authority: Lessons from England and New Zealand. *Phi Delta Kappan*, 78(8), 626-631.
- Wise, A. E., & Leibbrand, J. A. (2003). A professional model of accountability.

Wohlstetter, P., & Mohrman, S. A. (1993). *School-based management: Strategies for success*. New Brunswick, NJ: Rutgers University.

Wolfe, R., Childs, R., & Elgie, S. (2004). *Final report of the external evaluation of the EQAO's assessment process*. Toronto: OISE/University of Toronto.