

Reference: 20170044

31 March 2017

Thank you for your Official Information Act request, received on 15 February 2017.
You requested the following:

“A copy of any reports prepared by Dr Graham Scott and Helen Timperley on the school accountability system and any associated Treasury briefing/report documents.”

Information Being Released

Please find enclosed the following documents:

Item	Date	Document Description	Decision
1.	August 2016	Think Piece on Education Accountability – Helen Timperley	Release in Full
2.	31 August 2016	A systems perspective on accountabilities in the school system – Dr Graham Scott	Release in Full

Please note that although Dr Scott’s report is labelled ‘draft’, no further changes were made to the report and it can be considered to be final.

Please note that this letter (with your personal details removed) and enclosed documents may be published on the Treasury website.

This fully covers the information you requested. You have the right to ask the Ombudsman to investigate and review my decision.

Yours sincerely

Diana Cook
Acting Team Leader, Education and Skills

Information for Release

OIA 20170044

1.	Think Piece on Education Accountability – Prof Helen Timperley	1
2.	A Systems Perspective on Accountabilities in the School System – Dr Graham Scott	33

Think Piece on Education Accountability

Helen Timperley, PhD and Sarah Mayo, PhD

Faculty of Education

The University of Auckland

August 2016

RELEASED UNDER THE
OFFICIAL INFORMATION ACT

Acknowledgements: We would like to sincerely thank the people with whom we discussed the accountability issues at different layers of the New Zealand education system. At times they offered their comments with considerable passion. They were more informative in many ways than the written documents we consulted.

Contents

Introduction 3

What Do We Mean by Accountability? 3

Purpose of Accountability 4

Intersection between Accountability and Autonomy 5

The Characteristics of Effective Accountability Systems 6

Approaches to Accountability and the New Zealand Education System 7

 Student Outcomes 8

 Teacher and Leader Evaluation and Accountability 10

 School-level Evaluation and Accountability 11

 Accountability for Communities of Learners 12

 System Accountability and Coherence 13

A Closer Analysis 13

 Student Outcomes 13

 Professional Evaluation and Accountability 15

 Initial Teacher Education 17

 School Level Evaluation and Accountability 18

 Communities of Learners and Accountability 18

 System Coherence 23

Possible Re-settings 24

References 28

RELEASED UNDER THE OFFICIAL INFORMATION ACT

Introduction

This think piece, commissioned by The New Zealand Treasury, draws on national and international literature, as well as our knowledge of the New Zealand context, to consider the purposes of accountability systems and how the accountability settings in New Zealand's schooling system should be set in ways that drive sustained improvement in students' learning outcomes. It describes accountability levers at different levels of the system. Possibilities for change are considered in the context of New Zealand's education system in which schools have considerable autonomy. Attention is paid to the importance the Treasury places on accelerating Māori educational achievement.

We begin by defining what we mean by accountability, because it is an idea with many different interpretations but few working definitions. A brief section on the purposes of accountability systems follows: here there is much greater agreement among key authors, who conclude that improving outcomes for students, particularly those less well served by our education systems, is the primary function of accountability. Accountability systems do not sit in a policy vacuum, so we briefly consider the intersection between accountability and autonomy. There are a number of other unique features of the New Zealand education system to which we refer throughout subsequent sections of this paper as we describe accountability mechanisms at different levels of the system. Many of these features appear to reflect what is currently advocated to be effective in the international literature, yet New Zealand has persistent issues with equity in student achievement stratified along ethnic lines. For this reason, the section following this system description takes a closer look at how the accountability mechanisms play out at each level of our education system. The paper concludes with possible re-setting of accountability mechanisms throughout the education system.

As indicated above, the observations and ideas in this paper have been informed by a search of recent international literature on accountability in education. We have also examined research, documents and reports specific to the New Zealand context, including three randomly selected CoL plans on the Ministry of Education's website (<http://www.education.govt.nz/assets/Documents/Ministry/Investing-in-Educational-Success/Communities-of-Schools>). In addition, we have spoken to a range of people from different agencies together with appointed leaders of Communities of Learning (CoL), and principals in both English and Māori medium schools. As we searched through documents and held discussions with these educators we have uncovered unexpected levels of complexity and recent change, which have made the task of ensuring accuracy more difficult than anticipated. As an illustration of this complexity, when we asked a principal of a Māori medium kura to make a list of those to whom she was accountable, she declined with the comment, "*It would take me all day*". The sheer number of accountabilities and their complexity may be at the heart of the system problem. Given this complexity, and the constraints of funding and time for this thinkpiece, much more may need to be investigated before decisions are made about any redesign.

What do we mean by accountability?

In the papers accessed for this review, few authors define what they mean by accountability and different ideas clearly underpin their ideas. For example, Fullan et al. (2015) write, "Simply stated, accountability is taking responsibility for one's actions" (p.4). In contrast, Hargreaves defined accountability as "the remainder that is left when responsibility is subtracted" (Hargreaves, 2008, p.

37). The organisational literature emphasises the role of others in the accountability equation. For example, for Kogan (1986), accountability is “a condition under which a role holder renders an account to another so that a judgment may be made about the adequacy of the performance” (p. 25). Tetlock (1998) provides a more cognitive emphasis: “the implicit or explicit expectation that one may be called on to justify one’s beliefs, feelings and actions to others” (p. 255).

In the reviewed literature, each of the elements of these definitions is encapsulated in different ways and it is important to consider them together:

- Taking responsibility for one’s actions
- Providing an account to others about the adequacy of performance
- Justifying beliefs, feelings and actions to others.

Also pervasive in much of the reviewed literature is the idea of high stakes / low stakes accountability, but we have taken these to be the consequences of accountability once the account has been provided. High stakes consequences are those where explicit rewards or sanctions follow from providing the account. Low stakes consequences are those where no clear rewards or sanctions follow (Wößmann, Lüdemann, Schütz & West, 2007). It needs to be noted that perceptions of whether the stakes are high or low stakes may differ among individuals in similar situations because less tangible factors, such as the impact on professional identity, come into play.

We consider that this issue of different definitions and separating the accountability “processes” from the “stakes” is important in order to progress the discussion. Hattie (2015), for example, conflates the process with the stakes in his advocacy for the development of collaborative cultures in which teachers work to achieve “at least a year’s growth for a year’s input” (p. 5). The collaborative processes he describes include all the above three bullet points, including a strong emphasis on measurement of student achievement and progress when providing the account to others. In the same paper he writes, “there can be no whiff of accountability based on students’ test scores” (p. 24). Presumably he is referring to the possible negative consequences of blaming and shaming – a sentiment with which we strongly agree.

Purpose of accountability

There is little argument with Fullan et al.’s (2015) stated purpose for accountability: “At the core of accountability in educational systems is student learning” (p. 4). These authors emphasise that student learning is not just improvements in test scores but deeper and more meaningful learning for all students. In New Zealand, attention to our educational inequities must be at the heart of the purpose of our accountability mechanisms and systems.

Some educational jurisdictions, most notably in the United States, have enacted this purpose by focusing their accountability systems directly on student performance with high stakes consequences. However, this approach has justifiably been widely criticised because education is more complex than these simple input – outcome assumptions. There is considerable evidence that systems that directly and exclusively target student achievement have resulted in distortions, gaming of the system, and little real improvement. It has not been the approach of high performing or improving systems. In particular, it undermines the trust required to develop a learning system (Daly & Chrispeels, 2008). Tucker (2014) summarises the evidence provided by many others:

The result is very low teacher morale, plummeting applications to schools of education, the need to recruit too many of our teachers from the lowest levels of high school graduates, a testing regime that has narrowed the curriculum for millions of students to a handful of subjects and a very low level of aspiration. There is no evidence that it is contributing anything to improved student performance, much less the improved performance of the very low income and minority students for which it was in the first instance created. (p.2)

Internationally, there is at least an implicit improvement agenda in all educational accountability systems, even in the United States where it was hoped that a focus on student test scores would raise those scores (Lasky, Schaffer, & Hopkins, 2008). More recently, this improvement agenda has become more explicit as attention has turned to higher performing systems where the accountability architecture and processes are integrated into a wider improvement agenda. In these systems, evaluation and accountability focus not only on student performance, but also on the quality of instruction, school development and professional learning (Schleicher, 2015, p.46).

Intersection between Accountability and Autonomy

Accountability approaches do not sit outside other characteristics of the system but rather interact closely with them. New Zealand is noted for its high level of autonomy over curriculum, assessment and resources, features which are frequently identified with successful school systems (OECD, 2010). The benefits, however, are not always realised. Two authors (Farrar, 2015; Suggett, 2015) argue that these benefits accrue only when accountability is also rigorous. For example, Suggett concludes from an analysis of OECD reports (OECD, 2010; 2012) that:

The institutional context for autonomy matters, and accountability in particular makes a difference. Where accountability is weak, autonomy in both managing resources and determining curriculum and assessment can in fact worsen performance. When accountability is strong, autonomy is an advantage, although how much autonomy and where also makes a difference. (pp. 13-14)

Others argue for the importance of professional capacity for accountability to work in an environment of relative autonomy. Fullan et al. (2015), for example, insist that any attempt to reset accountability structures must begin by building the professional capital of teachers and leaders. By this they mean “the collective capacity of the profession and its responsibility for continuous improvement and for the success of all students” (2015, p. 6). Both Farrar (2015) and Greany (2015) additionally argue that leadership capability is essential to the accountability settings inherent in a self-improving system: “Greater autonomy and more responsibility, not only for the improvement of your own school but also for others, makes headship ... daunting” (Farrar, 2015, p.7).

The OECD report on 2012 PISA results found that the professional capital needed for autonomy to be effective in improving student outcomes could be compensated for and enhanced through collaborative practices (OECD, 2013a): this signals the interdependence of accountability, autonomy, professional capacity, and collaborative practices.

The Characteristics of Effective Accountability Systems

There has been increasing international interest in effective accountability mechanisms and systems. Part of the reason for this is a reaction to the United States' failed emphasis on measures of student achievement with high stakes consequences. Studies of effective schools (e.g. Elmore, 2004), clusters of schools (e.g. Farrar, 2015), and systems (e.g. Fullan et al., 2015) have led to advocacy for internal accountability based on collaborative cultures characterised by high trust, collegial commitment to collective responsibility for continuous improvement, and success for all students. Such systems develop rather than undermine teacher professionalism. Elmore et al. (2004), for example, identify that the main feature of successful schools was their collaborative cultures that combined individual responsibility, collective expectations, and corrective action.

Indeed, there is considerable evidence from New Zealand research that rapid gains can be made in achievement and can be sustained with the development of high quality internal accountability processes among leaders and teachers (Bishop et al., 2009; Phillips, McNaughton & MacDonald, 2001; Timperley & Wiseman, 2003; Timperley, Parr & Bertanees, 2009; Timperley & Parr, 2009). The processes have typically involved school-based professionals working with external experts to:

- Collaboratively examine evidence of student learning;
- Analyse the quality of their professional practices and how they might be contributing to the current student profiles of achievement;
- Develop and implement specific targeted pedagogical practices with integrity; and
- Check the impact on student learning through iterative cycles of inquiry until the most difficult to teach students have made demanding progress based on high expectations.

More recently the Education Review Office (2016) identified similar characteristics of effective self-review that resulted in accelerated outcomes for priority learners. In contrast, those schools whose students made little progress usually carried on offering more of the same.

Although the emphasis of the current international literature is on internal accountability and collective responsibility, typically around a range of soft measures, there is an indication that a mix of hard (typically external) and soft (typically internal) measures are needed. Fullan et al. (2015) describe internal accountability as when individuals and groups "willingly take on personal, professional and collective responsibility for continuous improvement and success for all students" (Fullan et al., 2015, p.4), while external accountability is "when system leaders reassure the public through transparency, monitoring and selective intervention that their system is performing in line with societal expectations and requirements" (Fullan et al., 2015, p.4). These authors claim that internal accountability must *precede* external accountability if lasting improvement in student achievement is the goal:

The solid and mounting evidence on the fundamental impact of internal accountability on the effectiveness and improvement of schools and school systems contrasts sharply with the scarce or null evidence that external accountability, by itself, or as the prime driver, can bring about lasting and sustained improvements in student and school performance. There is, indeed, a growing realization that external accountability is not an effective driver of school and system effectiveness. (Fullan et al., 2015, p.6)

One way to conceptualise accountability is in terms of hard or top down measures versus soft or collegial measures, as illustrated in Figure 1.

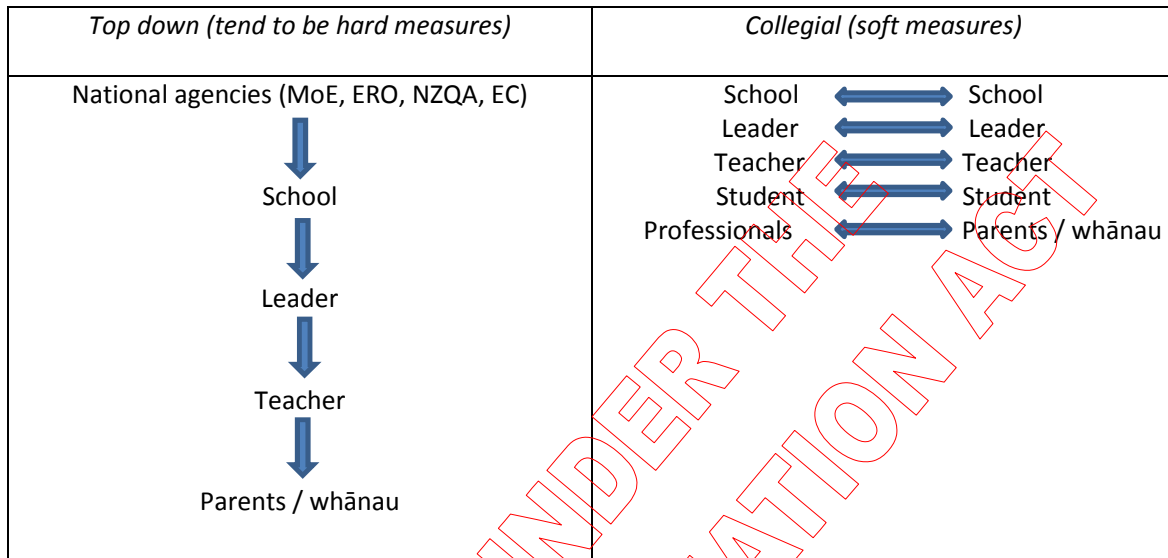


Figure 1. Top down and collegial accountability measures.

In addition, bottom-up or reciprocal accountability involves holding the system accountable for providing the necessary conditions for improvement, such as adequate resources and training or professional development, at the same time as holding schools and teachers accountable for performance and improvement (Fullan et al., 2015). Reciprocal accountability might also involve assessing student outcomes in terms of the context and conditions in which they have been achieved. The best example we know of reciprocal accountability in New Zealand occurred in the Literacy Professional Development Project that resulted in very high gains in literacy achievement (Timperley & Parr, 2009). Ministry of Education officials, responsible for the programme, adjusted the policy settings as research evidence emerged about aspects that were hindering success.

Approaches to Accountability and the New Zealand Education System

As noted in the previous section, collective responsibility for improvement has been identified repeatedly in the international literature as central to internal accountability processes having a positive impact on student outcomes. An important consideration in the New Zealand context is that all official bodies in New Zealand, including the teacher unions, have an explicit commitment to the Treaty of Waitangi and the obligations for equity this carries for Māori learners, typically through reference to *Ka Hikitea* (Ministry of Education, 2013). There is also official collective commitment to improving the achievement of Pasifika learners.

In addition, current approaches to accountability, and the processes surrounding them, include a mix of hard and soft measures, largely identified as effective in the international literature and described in the previous section. They are designed to enhance teacher professionalism, collective responsibility and school improvement. Given this situation, the difficult question to answer is why these measures, in their present form and application, have not achieved the proposed gains in

student achievement, particularly for those learners who are not well served by our schooling system. In order to answer this question we briefly describe the accountabilities at student, teacher, leader, school, CoL and national level. A subsequent section takes a closer look at why these accountabilities may not be achieving their intended purpose.

Learner Outcomes

Evidence of outcomes for learners underpins all other accountabilities within most education systems (OECD, 2013b), as they do in New Zealand. This serves to put learners at the centre. Although focusing on learner achievement to the exclusion of other measures typically has a negative impact (particularly when associated with high stakes), student learning must be in the mix because it is the central purpose of schooling.

The student assessment system in New Zealand has a strong formative emphasis for both students and their teachers (Ministry of Education, 2011a). In primary schools, National Standards depend on teacher expertise, with their emphasis on overall teacher judgment, and are designed to promote teacher professionalism with the accompanying rich materials to assist them to make those judgements (<http://assessment.tki.org.nz/>). The Progress and Consistency Tool (PaCT) is designed to bring consistency to teacher overall judgements while providing information on the next steps in a learner's progress.

At upper secondary school, NCEA is a standards-based system and has more characteristics consistent with countries that perform significantly above average with high equity on PISA (OECD, 2012) than other systems. A recent analysis of upper secondary qualifications by Daniell (unpublished¹) identified the following characteristics of qualifications in the high achieving high equity quadrant of PISA, all of which apply to NCEA:

- Good quality pathways without dead ends
- Effective links to the world of work
- Attractive not just to an academically-included elite
- All can complete the equivalent of upper secondary studies/qualifications
- All have the opportunity to pursue tertiary studies if they wish
- No system-level policies associated with school and student failure, e.g. grade repetition, early streaming
- Opportunity to attain high level skills, regardless of personal and socio-economic circumstances.

The flexibility of NCEA led to some concern that the harder accountability measure of the BPS goal of 85% of school leavers achieving NCEA Level 2 was being met by increasing internal and vocational standards at the expense of traditional academic standards. An analysis of the shifts in credit composition in 31 decile 1-4 schools and 8 decile 5+ schools between 2009 and 2014 does not appear to bear this out for any group of students, including those of Māori and Pasifika ethnicities (Hynds & Courtney², unpublished). The four categories used in Figure Two include 'traditional-discipline' in 'arts' and 'science' that have a full complement of achievement standards, 'vocational' courses that are based more on unit standards, and 'contextually-focused' courses that offer a

¹ This analysis was undertaken as a study in a PhD thesis that is not yet completed.

² Hynds, A., & Courtney, M. unpublished data 2009 to 2014 shifts in credits obtained in NCEA Level 2

reduced number of credits, predominantly from internally assessed unit standards with “closer links to students’ everyday life contexts of future work or leisure”. Over the 1,217,548 credits, the proportional shifts in these schools in all categories were negligible (Science $d=0.00$; and arts $d=0.04$; contextual $d=0.03$; vocational $d=-0.08$).

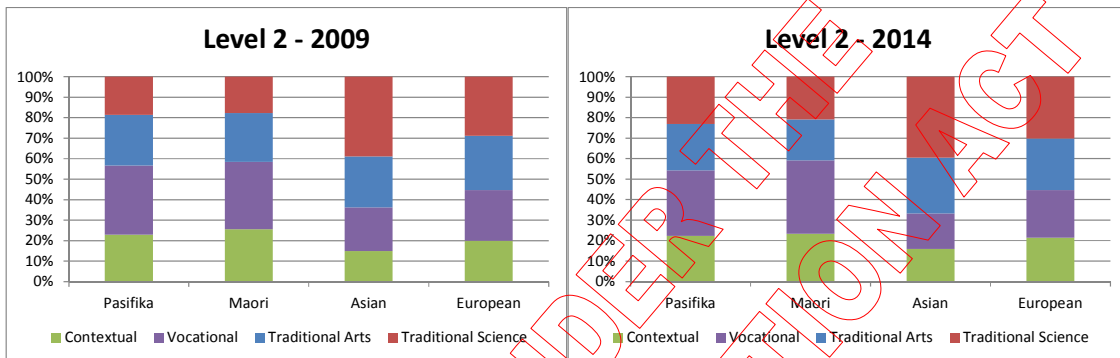


Figure 2. Shifts in the proportions of NCEA credits in each of four categories

Nor did the proportion of internal and external credits attained by Māori students change significantly between 2009 and 2014 in these schools ($d=-0.05$; see Figure 3), although it needs to be noted that a lower percentage of these students in these predominantly lower decile schools in the sample gained their qualification through internal credits than Māori students nationally.

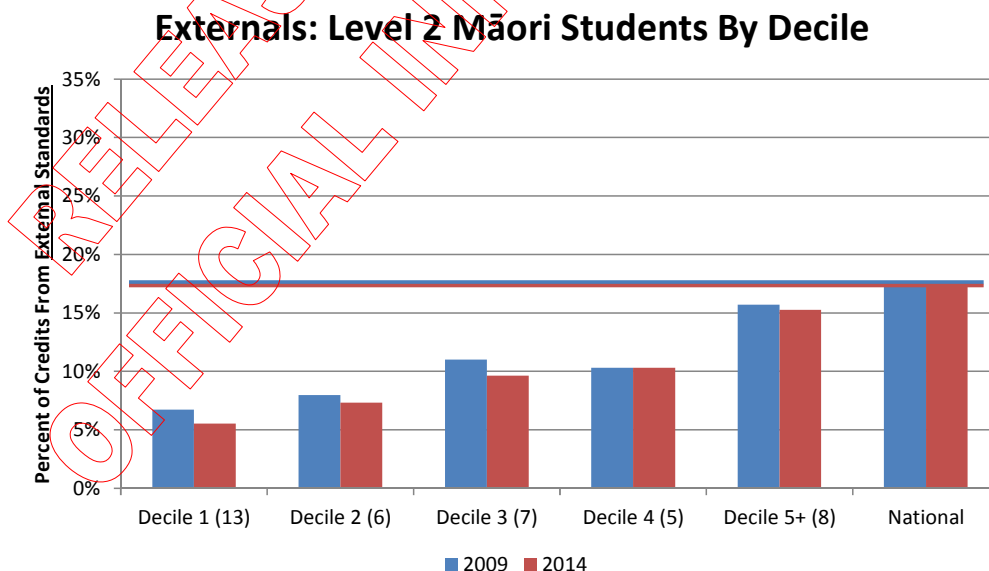


Figure 3. Shifts in the proportion of external credits between 2009 and 2014 for Māori students

Although the pattern of shifts in external credits was more complex across deciles for Pasifika students (see Figure 4), the overall effect size was similar to Māori ($d=-0.06$). Interestingly, Students who identify as European exhibited the largest shift away from level 2 external examinations of any group ($d=-0.13$)

Externals: Level 2 Pasifika Students By Decile

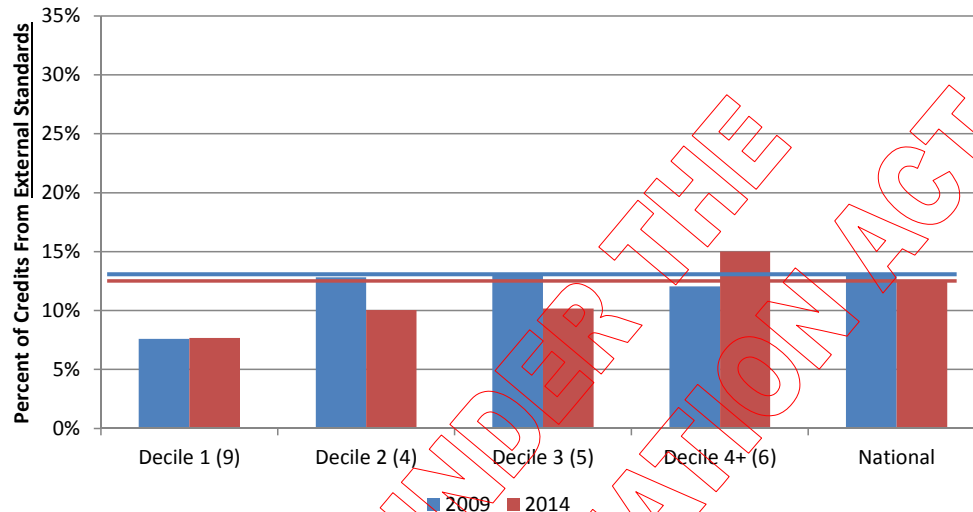


Figure 4. Shifts in the proportion of external credits between 2009 and 2014 for Pasifika students

The BPS NCEA target appears to have had the desired effect of improving attainment for Māori and Pasifika students without evidence of a strong distorting effect on the type of credits undertaken, but has done little to impact on the relativities in achievement for Māori and Pasifika learners. In order to really make a difference to equity issues, an ongoing focus needs to continue on the 15% continuing to not achieve NCEA level 2.

Teacher and Leader Evaluation and Accountability

The emphasis of teacher appraisal in New Zealand is developmental with inquiry into practice incorporated into the *New Zealand Curriculum* (Ministry of Education, 2015b). Graduating teachers must demonstrate they meet specific standards

(<https://educationcouncil.org.nz/content/graduating-teacher-standards>) set by the Education Council of Aotearoa New Zealand (Education Council) that provide guidance for expectations for competence. The Education Council is also responsible for accrediting initial teacher education programmes.

Teacher performance is also subject to an annual process of attestation by their principals against professional standards negotiated as part of separate union collective agreements with the Ministry of Education. Every three years teachers must renew their practising certificate. Part of the process involves principals endorsing teachers' performance using the Practising Teacher Criteria formulated by the Education Council to ensure teaching practice meets expectations. Teachers who do not meet these criteria after appropriate support are referred to the Education Council competence authority.

School principals are appraised on a set of professional leadership standards with attestation for performance by principals' appraisers and signed off by the Chair of the Board of Trustees. When it comes to the three yearly renewal of their practising certificate, appraisers are also required to attest to performance using the Education Council's Practising Teacher Criteria.

For both teachers and principals, the Education Council also recommends the inclusion of criteria developed from *Tātaiako* in an effort to heighten the importance of cultural responsiveness. Nga kura a iwi tumaki have developed a document with the Ministry of Education, “Tu Rangatira”, which is used on a voluntary basis. In addition, current guidance provided by the Education Council is to integrate both accountability requirements and developmental needs through an appraisal cycle.

While not an accountability mechanism for principals in the strict sense of the concept, a strong incentive for the way in which principals work is probably the relationship between the size of the school roll and principals’ salaries.

School-level Evaluation and Accountability

There is no shortage of bodies to which New Zealand schools are accountable, including the Ministry of Education, the Education Review Office (ERO), the New Zealand Qualifications Authority (NZQA) together with each school’s Board of Trustees, local community, and increasingly to iwi.

Under the Education Act 1989 all schools are expected to be involved in an ongoing, cyclical process of evaluation for improvement. Through the annual reporting process they are required to report on the achievement of their students, their priorities for improvement, and the actions they plan to take. Currently, the Ministry of Education requires schools to submit charters together with annual plans and targets and an analysis of variance from the previous year’s targets. These documents become the agreement between schools and the Ministry. The recent change to simplifying the process that essentially involves a letter of acknowledgement from the Ministry rather than detailed comment has been welcomed by the sector although there appears to be considerable regional variation in this practice.

The Education Review Office (ERO) was set up to evaluate and report publicly on the quality of education provided in New Zealand Schools and early childhood education services, and on the effective use of public funds. Where ERO determines that students’ learning or welfare is at risk, intervention by the Ministry of Education is recommended.

Since its establishment, ERO’s approach to external evaluation has changed from a compliance/accountability orientation to an accountability/improvement orientation. Since 2003 ERO has published school evaluation indicators for use in both internal and external evaluation. The complementary evaluation approach (Mutch, 2013) in 2008 was intended to strengthen the relationship between evaluation and improvement. The differentiated review cycles allowed evaluation resources to be more deliberately focused on education institutions that were not performing well.

In 2016, in response to the ongoing challenge of inequitable outcomes for particular groups of students in New Zealand schools, ERO introduced the *Accelerating Māori Student Achievement* approach in the external evaluation of primary schools. The main evaluative question is: *How effectively does this school respond to Māori students whose learning and achievement need acceleration?* This approach combines a sharpened accountability focus on learners who are not achieving equitable outcomes with a focus on improvement through building internal evaluation capacity (<http://www.ero.govt.nz/how-ero-reviews/accelerating-student-achievement-maori/>). The *School Evaluation Indicators* used by schools and ERO have also been revised in July of this year (ERO, 2016).

School principals report to their Boards of Trustees monthly and to their communities annually in ways that, in theory, make their students' overall achievement and other important indicators public. Primary schools are also required to report the numbers of students who are at, above or below national standards in reading, writing and mathematics to the Ministry of Education. These reports are also publicly available. This feature of using overall teacher judgement as the basis for public reports is unusual internationally, where public reporting is more typically in the form of results of national tests (Tucker, 2014).

Secondary schools are also accountable to NZQA for the quality of the NCEA standards and moderation processes. Aggregate NCEA results are also publicly available.

Accountability for Communities of Learning

In a systematic examination of improving education systems, Mourshed, Chijioko and Barber (2010) found that a middle tier or mediating layer, either pre-existing in the form of a district or local authority, or specially created in the form of a school cluster or community of learning, was a consistent and vital feature of all the successful systems they studied. This middle layer serves the functions of providing targeted support, acting as a buffer between the school and the centre, and serving as a channel to share and integrate improvements across schools (Mourshed et al., 2010). Since the initial analysis (Mourshed et al., 2010), there has been increasing pressure to position New Zealand's schools within more collective systems (Wylie, 2012, p.3; OECD, 2015b).

The recently developed Communities of Learning (CoL), where schools work together under the guidance of designated lead principals and teachers, are designed to fill the gap in the 'middle layer' between central government and schools in the New Zealand education system. This idea has been promoted and acclaimed by a number of high profile international experts. Munby and Fullan, for example, express the belief that "all schools, 100%, should be involved in focused, productive networks within which leaders, teachers and students challenge, support, innovate and learn from one another in ways that measurably improve outcomes" (Munby & Fullan, 2016, p. 5). These authors predict that all schools in New Zealand will be in a CoL by the end of 2017. Similarly, Hargreaves (2012) emphasises the importance of creating roles for high-performing head teachers, senior leaders and teachers to act as system leaders, as has been done in CoL.

The vision for CoL is to promote a learner-focused system through improved educational pathways. Accountabilities to achieve this vision are intended to be primarily collegial and promoted through professional collaboration, as indicated by the two references to accountability in the *Community of learners: Guide for schools and kura* (Ministry of Education, 2016). According to this guide, the appointment of the lead roles "will create the foundations for collective responsibility and accountability across your Community of Learning" (p. 3). On the following page, reference is made to collegial accountability: "The hallmark of your mutual commitment will be sharing responsibility and accountability for the outcomes of all the akonga/students in your Community of Learning" (p. 4). CoL are also expected to put in place "stewardship, governance and management structures" (Ministry of Education, 2016, p.3).

The architecture of CoL is also designed to promote a learner-focused system in which schools from ECE, primary and secondary sectors work together to create learner-designed pathways. In many ways CoL are designed to be consistent with Elmore's (2004) description for improvement as

“change with direction, sustained over time, that moves entire systems, raising the average level of quality and performance while at the same time decreasing the variation among units, and engaging people in analysis of and understanding of why some actions seem to work and others don’t” (p. 57).

System Accountability and Coherence

At the system level, a number of agencies are responsible for school level accountability. They include the Ministry of Education (MoE), the Education Review Office (ERO), the New Zealand Qualifications Authority (NZQA), and the Education Council of Aotearoa New Zealand (Education Council). All these agencies prioritise a collegial and developmental orientation to their work with schools, with a focus on building capacity rather than on punitive accountability systems. Their accountabilities have considerable overlap and are too complex to identify specifically and separately.

Early Childhood Education Services

Although the brief for this paper focused on schooling, recent government policy has framed the education system as 0-18 years. It is surprising, therefore, that this sector has been excluded from this brief.

A Closer Analysis

In New Zealand the rhetoric of commitment to equity and improvement-oriented professional accountabilities have not achieved the outcomes to which the education system aspires. OECD comparisons demonstrate declining levels of achievement at age 15 years, one of the closest relationships between socio-economic status and student achievement, and high within-school variability in learner outcomes (OECD, 2016). Therefore, it is important to take a closer look at the above description with the specific lens of searching for an explanation about why our approaches to accountability, apparently so consistent with effective systems, have failed to achieve desired outcomes. In this section, we will examine potential issues at each system level, paralleling the description above. We have based possible re-settings in the final section of this paper on this analysis.

Student Outcomes

As noted above, National Standards form the data used at a system level of primary schools. While the approach to the development and assessment of National Standards using overall teacher judgements (OTJs) is consistent with the research literature on high quality formative assessment (Ministry of Education, 2011a), this approach has presented some specific challenges. OTJs require teachers to consider a variety of evidence to decide how well an individual learner is progressing in relation to a given standard. This task is complex and has led to inconsistencies (Ell et al., 2015).

A recent report on the *National Standards: School Sample Monitoring and Evaluation Project, 2010-2014* (Ward & Thomas, 2016) concludes from evidence collected from approximately 100 schools that:

Considered together, this body of evidence strongly suggests that OTJs lack dependability, which is problematic as OTJs are a central element of the National Standards system. It

should be noted that there is no suggestion that all OTJs are inaccurate, but evidence indicates that a reasonable proportion may be. (p.2)

This report also reports an improvement in achievement over time but cautions:

Given the magnitude of the improvements in achievement that are suggested by the OTJ data, the evidence that suggests OTJs lack dependability, and evidence about patterns of student achievement in New Zealand from international studies, the OTJ data cannot be taken as evidence that student achievement is improving over time. (p.3)

The development and introduction of the Progress and Consistency Tool (PaCT) by the Ministry of Education has sought to address this inconsistency. The reluctance by many primary schools to use this tool appears to suggest that they do not value high quality evidence sufficiently to ensure the accuracy of OTJs. Some of those we spoke to in preparation for this paper indicated that their reluctance was more politically motivated than educational. Whatever the reason, evidence from OTJs is fundamental to understanding individual learner, class, school and system performance, and forms the touchstone for other levels of the accountability system. This issue of unreliability, therefore, is of serious concern. Suggestions about how to establish these is beyond the scope of this paper but, for the reasons cited earlier in this paper about the lack of evidence about the effectiveness of high stakes, test-based accountability in the United States, we would not recommend that the New Zealand system adopts national testing. The solution needs to be found in improving the quality and reliability of teacher judgements.

This problem of inconsistency is exacerbated by increasing the stakes for these assessments through comparative publication of results across schools. As noted earlier, increasing the stakes related to student achievement can lead to adverse consequences of distortion, focusing on what is tested, and narrowing the curriculum. Publishing OTJs, with an unknown proportion of these judgements essentially unmoderated and inaccurate, will almost inevitably lead to inflated and unreliable judgements.

A second issue is one of potential bias in the teacher judgements. A recent paper by Blank, Houkamau and Kingi (2016) identified the possibility of negative cognitive bias in teachers' judgements about the achievement of Māori students. This bias has been more clearly documented in a study comparing students' achievement on standardised tests and teachers' overall judgements of achievement in National Standards (Meissel, Meyer, Yao & Rubie-Davies, in review). This bias is evident for Māori and Pasifika students, particularly for boys in writing. The concern is that students from Māori and Pasifika ethnicities are not only perceived by their teachers to achieve at lower standards than those from other ethnic groups given the same standardised achievement score, but also that this perception may be reflected in their expectations of these students and the opportunities to learn that are provided (Turner, Rubie-Davies & Webber, 2015).

In secondary schools, NCEA is the most frequently used student-level accountability measure, with no formal reporting requirements prior to Year 11. The variability in outcomes in NCEA Level Two by schools of the same decile is already documented by The Treasury. This variability continues at NCEA Level 3 and University Entrance (see Figure 5). Although we have included this analysis in the section on student outcomes, the accountability question does not rest with the students. Rather it rests on the relevance and quality of learning offered to them. So the central issue that these data highlight

is, “Who holds schools to account for this variability in performance?” It is unlikely that individual Boards of Trustees or the community have access to this kind of data. It needs to happen at a system level.

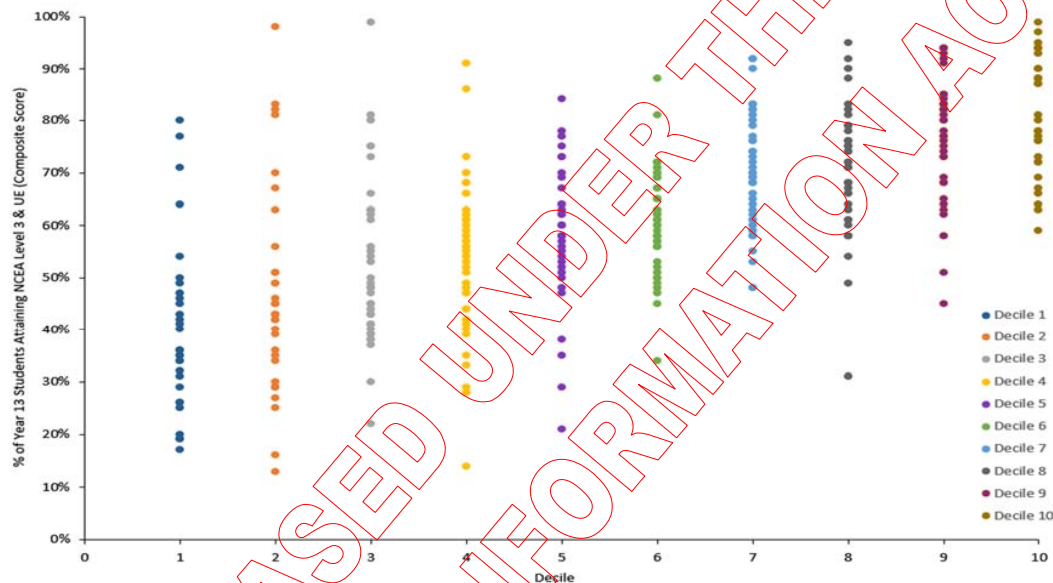


Figure 5. Within school variability by decile for NCEA level and UE aggregate score (Source Daniell, and Stephens, unpublished³)

Professional Evaluation and Accountability

Accountability measures and mechanisms for teachers are complex. We have already commented on the distortions and problems that occur when student achievement results are used to the exclusion of other criteria. Problems also arise with observations of practice. A large scale study in the United States (Measures of Effective Teaching [MET] Project, Bill and Melinda Gates Foundation) identified that single observations with a single observer are typically unreliable in terms of evaluation. The most reliable evaluations comprised a mix of multiple observations of teaching practice by more than one observer using clear criteria together with student feedback and measures of student progress. This is a relatively time consuming mix to administer but if it is based on informed judgement and approached from a developmental perspective with adequate feedback, an adapted version could prove useful. It is similar from the approach by Bishop, Berryman, and Cavanagh (2009) in Te Kōtahitanga, with significant shifts obtained for Māori students in secondary schools.

Internationally, teaching and leadership standards (e.g. AITSL, www.aitsl.edu.au) are used increasingly as the criteria against which teacher practice is judged – serving both an accountability and developmental function as the criteria become more demanding. In New Zealand, as described

³ These data were analysed as a sampling strategy for a PhD thesis on NCEA pathways.

above, there are multiple competing standards and criteria with little coherence between them, an issue noted in the OECD report on assessment and evaluation (Nusche et al., 2012). On the ground, one principal described the different systems as “*a nightmare*”. The focus is on multiple reports, placing a high administrative burden on school leaders and teachers for little apparent improvement gain. The Education Council is currently working with the sector to bring the professional standards and practising teacher criteria together into a simpler, more coherent set of standards which, if achieved, will at least provide the basis for more transparent and reliably assessed accountability.

At least up until recently, the accountability measure of attestation of teachers’ competence by their principals has been weak. A country report for the OECD (2010) on New Zealand’s approach to assessment and evaluation stated that no teacher had failed attestation against the professional standards. While in practice this is probably not an accurate depiction of the reality, it is indicative of the weakness of attestation as an accountability mechanism.

The Practising Teacher Criteria, used as a benchmark for the full registration and re-registration of teachers, together with compulsory reporting, has resulted in a small but growing number of teachers brought to the Education Council’s attention.

Two recent publications indicate that, despite the efforts of the Education Council and some schools, appraisal may not be serving to focus the profession on outcomes for Māori. Hynds, Averill, Hindle and Meyer (in press) found little evidence in their 75 in-class observations of culturally responsive practices in the nine participating secondary schools. Early indications from an audit of teacher appraisal practices undertaken by ERO for the Education Council has identified highly variable appraisal practices in schools.

Making accurate judgements, however, is only one part of the puzzle if they are to result in improvement. The other part is about knowing how to use evidence to decide what to do once the judgements are made at each level of the system (student, teacher, schools, nationally). There are several indications that this capability is weak in many schools. We will illustrate difficulties around this issue by drawing on the final report of the Starpath Phase Two Project (Kiro et al., 2016). A matched school comparison showed that “the matched schools outperformed, or performed at the same levels, as the Starpath schools” (p. 19) despite considerable resources and professional support available to the Starpath schools. The report noted that the different interventions associated with Starpath were variably implemented across schools and teachers, even when the evidence supported the efficacy of particular practices over others in terms of learner outcomes. With respect to literacy, for example, the report notes that:

the higher the quality of literacy instruction (before or after the intervention) the higher the pass rates for SLAs [Literacy achievement standards] The estimate from the statistical modelling is that an average increase of one additional aspect of literacy instruction after the intervention would result in increased odds of passing English, mathematics and science SLAS standards It is important to note that the intervention did not produce uniform shifts in literacy teaching practices and this did not translate into generally higher pass rates for students in high literacy standards. (p. 17)

We are not intending any specific attribution about the causes underlying this finding because it may be a system issue (particularly related to multiple interventions in these schools) as much as a

teaching issue or some combination. Possible causes may also include Starpath's approach to improvement in which literacy intervention was only one of a number of processes of change, and the complexity of schooling in low decile schools. Rather we consider this is an issue in need of serious investigation.

Similar statements are made in the Starpath Final Evaluation Report (Kiro et al.,2016) about improvement in school leadership and the issue of reducing within-school variability (a central accountability issue) despite a major focus in this area. The report concludes:

There was no improvement in measures of leadership cohesiveness average ratings between senior and middle leaders were significantly different on the core issue of addressing variability of teaching in schools. (p. 6)

The report goes on to note that,

A period of intensive intervention in five of the schools towards the end of Phase 2 served to highlight the significance of the multiplicity and complexity of the issues faced by school leaders on a daily basis. (p. 6)

One important strand of work in the Starpath intervention was a focus on goal setting for school leaders - a process that is central to accountability and to the new work of CoL. An analysis of school plans revealed that,

There were too many goals and targets related to student retention, engagement and achievement. School leaders struggled to recall their school goals; and indeed it was also challenging for the leadership facilitation team to identify many current priorities from school documentation. While, on average, the analysis team found four academic goals, most schools had large numbers of associated targets (on average nine), as well as a number of separate interventions. These were often additional programmes. (p. 17)

We will return to this issue in our discussion of accountabilities for CoL.

Initial Teacher Education

The initial teacher education space is complex in a different way. While the Education Council has primary responsibility for accrediting programmes, the plethora of providers and programmes makes rigour in this process very difficult. Twenty-five different providers and 156 programmes (Education Council, 2016, p.7), each with different entry criteria, philosophical approaches and configurations of courses that must be evaluated for accreditation purposes, make rigour around accountabilities very difficult.

There is also an issue of the number of government agencies in the initial teacher education space. From a programme perspective, the multiple accountabilities mirror those for schools. For example, the newly introduced Masters of Teaching programmes are currently being monitored and evaluated separately by the Education Council, Ministry of Education officials, and externally contracted evaluators. The University of Auckland Post Graduate Diploma in Teaching (Secondary Field-based) – commonly referred to as the Teach First programme – is being evaluated for the Ministry of Education by the New Zealand Council for Educational Research. All programmes are externally monitored annually or biennially by the Education Council of New Zealand, and are

required to undertake an ECNZ panel re-approval every six years. Programme providers are also encouraged to have their own internal formative evaluation but are not required to demonstrate impact. In addition, programme leaders are required to complete a graduating year review for the Committee on University Academic Programmes, normally within 3 years of the graduation of the first cohort of students from a programme. None of these agencies has sufficient time or funding to investigate the impact of these programmes with rigour. We suggest that these multiple reporting requirements are more likely to detract from rather than add to programme quality as leaders prepare different sets of evidence to meet different agency requirements using different criteria for evaluation then take the time to meet with each one to explain what is happening. As one initial education leader said with some frustration, *“It’s hard to keep focused on what is important”*.

The sometimes suggested solution to the variability of outcomes from initial teacher education of locating more of the preparation in schools is unlikely to improve rigour or reliability, given that schools’ and teachers’ prime responsibilities are for student learning and the even more diffused accountabilities that would be associated with locating teacher education in potentially hundreds of schools. In fact, it may be that the opposite solution is likely generate more robust and reliable outcomes – namely, reducing the number of providers (from the current 25) and monitoring them through a clearer system of accountability.

Requirements for teacher education candidates to demonstrate that they have met the graduating standards are variable: course requirements and assessments are included in the monitoring evidence, but the quality of graduates exiting programmes into the profession is not monitored beyond statements they have met the graduating standards, often in the form of portfolios that are variably moderated.

School Level Evaluation and Accountability

As noted above, there is no shortage of agencies to which schools report in relation to a range of statutory accountabilities: the Ministry of Education, ERO, NZQA, individual school Boards of Trustees, local communities and increasingly to iwi. Both ERO and the Ministry have supported an improvement-oriented approach to internal evaluation that intersects with these agencies’ external review processes. The integration of internal and external evaluation is consistent with international trends towards schools taking responsibility for their own improvement (Farrar, 2015). While this is clearly a desirable outcome, the number of reports that schools must make to different agencies brings into question the balance between reporting and the time and effort needed to use these multiple reports for developmental purposes. This point was made by the leader of the kura when she declined to list all those to whom she reported on the basis it would take her all day. The sheer number of reports and the different frameworks currently being used by different agencies inevitably diminishes the time and effort that can be put into acting on the information in the report for improvement functions.

One of the two major agencies to which school report is the Education Review Office. The office has made several changes to their systems that are intended to enable more evaluative reporting, an increased focus on learners who are not achieving equitable outcomes, and follow up action together with internal evaluation capability building. However, the impact on current learner outcomes more accurately reflects the situation prior to these changes, so we have paid attention to them in this closer look because they highlight some systemic difficulties in this accountability space.

Until recently, ERO took a complementary evaluation approach (Mutch, 2013), which in reality led to primarily descriptive rather than evaluative reports. The public records were bland and did not adequately differentiate high performing from low performing schools, with the best indicator being the number of years before the next review. A survey of school principals by Wylie and Bonne (2014) highlighted the impact of this approach. These principals expressed a lack of confidence in ERO's evaluation judgments about schools other than their own and only 23% thought that reviewer judgments were consistent across schools.

Another difficulty is that the approach assumed the capability both to make evaluative judgements and to use the evidence to promote development and improvement. These are complex and skilled processes. While there is considerable evidence that schools that are effective in their internal evaluation processes are self-improving, there is little empirical evidence to support the efficacy of internal review as an improvement mechanism for all schools, particularly those performing in the lower bands (OECD, 2014). ERO's national evaluation reports since 2003, and more recently case studies of schools where internal evaluation promotes improvement in student outcomes, have identified the need for improved data literacy: "posing focused questions, using relevant data, clarifying purpose(s), recognising sound and unsound evidence, developing understanding of statistical and measurement concepts, and engaging in thoughtful interpretation and evidence-informed conversations" (Education Review Office, 2016, p. 41). These reports have highlighted the tension between schools being accountable for ongoing internal evaluation, but not having the evaluative capability to do so. There has been limited investment in a systematic approach to building this kind of capability across the system. This is a recurring theme in the New Zealand education system. High quality frameworks informed by the international literature are put in place without the capability-building piece of the puzzle being given adequate attention. The discussion of the development of CoL in the next section also underlines this point.

The previous version of ERO's school evaluation indicators, *Evaluation Indicators for School Reviews* (2010), did not help develop schools' evaluative capability. Two hundred and seventy eight indicators, of varying significance, were not organised around an underpinning theory of action (Earl, 2014). Nor did they reflect the key elements of effectiveness and improvement and their interconnectedness. Earl points out why indicators without a clear framework may not serve a strong evaluation function:

Although the indicators are identified with particular dimensions, there is no overarching theory of change to provide structure and direction for the process. Without such a theory, it is likely that schools will attend to some of the indicators in isolation, often with superficial treatment and results, rather than working from a theory of change that coordinates the work and guides the collaborative process of gathering, considering and interpreting evidence in the service of making changes in practice. (p. 16)

If accountability is to serve a developmental function, there remains an additional potential issue with follow-up for those schools whose performance is not judged to be adequate. Wylie (2012) notes that at any one time, ERO identifies 16-20% of schools as struggling to meet their responsibilities, and that this percentage has not changed over time. There does not appear to have been effective system approaches to working with these schools and this is needed if the country's equity problems are to be addressed. During the first author's time on the Professional Learning and

Development Advisory Group, for example, multiple instances were cited of schools in difficulty having several different Ministry-funded providers working in the same school, often with different intervention approaches. During the re-design process, one Māori medium PLD provider described how one small kura had three different professional development providers offering a total of six programmes. The Starpath Phase Two report (Kiro et al., 2016) similarly noted the difficulty the team experienced in the number of initiatives being offered in schools, thus competing with others for time and attention from school leaders.

Both international and New Zealand research on schooling improvement (e.g. Bryk et al., 2015; Timperley et al., 2010) conclusively demonstrates that single focus interventions have far greater impact on student outcomes than multiple interventions, particularly when these interventions are disconnected from one another. The redesign of nationally-funded professional learning and development may assist with this problem of duplication, but how effective it will be is dependent on the quality of decision-making across the regional Ministry of Education offices and panels, including the extent to which they hold those receiving funding to account for improvement.

Communities of Learning and Accountability

The introduction and development of CoL is another example of a potentially good idea, well informed by the international literature, but there is scope for more consideration of the capabilities needed at all levels of the system to make a real difference to the learning of those students least well-served by our education system.

The early rationale for establishing CoL (initially known as Communities of Schools) originated with a report on raising the status of the teaching profession. The additional funding was promoted as creating new possibilities for career pathways for excellent teachers who wanted to stay teaching but could not receive salary increments unless they moved into management positions (Bendikson, 2015). Some funding was also directed to new leadership positions. Over time, this purpose evolved into having a more direct focus on schooling improvement across the ECE and schooling sectors, with the funding for the new positions dependent on an approved plan with achievement challenges and associated targets. Accountability mechanisms are primarily focused on professional accountability, consisting primarily of soft measures with an emphasis on collaboration and collective commitment to continuous learning.

Achieving the vision for CoL as promoting a learner-focused system through educational pathways is complex and challenging. Given the focus of this thinkpiece on accountability, we have restricted our comments to the ways in which the accountabilities designed as part of the CoL developments are likely to promote this vision. We acknowledge the limitations of restricting this focus because accountabilities form only part of the context in which CoL are developed and operate. However, Fullan et al. (2015) argue that in such developments, accountabilities need to be explicit at the policy level.

One focus of accountability for CoL presumably relates to meeting the identified achievement targets in approved CoL plans. Each CoL is expected to identify 3-5 achievement challenges and a number of targets to be achieved within a 2-3 year time frame. The shape of the challenges and targets are monitored by the Ministry of Education and approved by the Minister's office. The process of achieving the targets appears to rest on collegially focused professional accountability.

Farrar (2015) has written most extensively on building “accountability rich cultures” in an essentially competitive system in a study of school clusters in England and Australia. She identifies the importance of a collective commitment to inquiry, learning and growth: “the approach, which is capable of building ‘trust-based and shared accountability’ within a competitive market-led system, has the potential to create truly sustainable and improving school-led systems” (p.3).

She also cautions that,

We all know of many cluster-based initiatives that have not achieved this promise – where impact has been weak, where any gains made have not been sustainable, and where substantial investment has been provided but it is difficult to know exactly what difference it has made. (p.8)

A common problem identified by Farrar was that accountability for achieving agreed goals was overridden by concerns about maintaining relationships and warns that,

a focus on maintaining the good relationships that they currently have can mean they are resistant to holding each other to account for outcomes. These clusters are in danger of becoming complacent; they can be deluded as to how good they really are. (p. 11)

Munby and Fullan (2016), strong advocates of the development of CoL, assert that the best way to increase collective accountability is through “the sheer power of effective peer review” (p.7). They also acknowledge that “if we are to move to a joined-up, fully networked system, what is being asked of relatively autonomous school leaders is a very significant cultural change” (2016, p. 10). These authors, together with others (e.g. Farrar, 2015) acknowledge the challenges in developing this kind of cultural change, particularly moving from individual responsibility to collective responsibility for changing students’ experiences of schooling in order to improve outcomes. Collective responsibility develops through closely examining the experiences of students and unpacking how the adults in the situation are contributing to the current patterns of achievement (Bryk et al., 2015; Timperley, Kaser & Halbert, 2014). The Treasury⁴ picks up this issue in multiple system level circles by identifying the need to answer the question, “Why is it so?”

The analysis from which the achievement challenges and targets are developed by the CoL is intended to form the basis for identifying what needs to change, particularly in the instructional core (Elmore, 2004). To be successful, persistence is required in relation to a single challenge, to which there is a high level of commitment until measurable changes have been met, before attempting to move on. Too many challenges at any one time typically result in limited improvement (Timperley et al., 2010; Bryk & Schneider, 2002; Bendikson, 2015).

Two questions arise from this literature in relation to the requirements for CoL to identify targets and achieve them through collaborative peer-to-peer accountability. The first is, ‘Are the current accountabilities likely to foster the kind of cultural change needed to improve student outcomes, particularly those not currently achieving as well as others?’, with the second, ‘Who, in this essentially voluntary system, is accountable for what?’

⁴ The Treasury (undated). *Higher Living Standards and Education*.

To answer this first question, we accessed a recent survey undertaken by the New Zealand Principals' Federation about the reasons principals joined CoL: only 20% stated that they were motivated to improve learning outcomes for students. We also examined the three CoL plans we had accessed. A focus on collective responsibility for changing the instructional core was not particularly evident in the plans. For example, one answered the question about "Why is this so?" by locating the achievement challenges in perceived deficits of students rather than the educational provision offered to these students. To quote from this plan:

The focus of our achievement challenges is literacy and numeracy, equitably attained by all learners with a specific focus on Māori, as essential for accessing current and future learning opportunities. Underpinning these challenges, it is our collective understanding that the root cause of underachievement is fundamentally attributable to the lack of key capabilities in students which are essential both now and in their future lives. As the NZC makes explicit, the capacity to manage oneself, relate to others effectively and participate and contribute, for example, has a direct impact on student learning and achievement.⁵

The other two plans largely by-passed asking or answering the question, "Why is this so?" Indeed, guidelines for CoL published by the Ministry of Education (2016) do not suggest they do so.

An additional problem evident in the plans was setting targets in terms of the percentage of learners reaching a specified level or a standard (e.g. 80% of students will ...). Given that the students of greatest concern are likely to be in the untargeted 20% percent, this form of target can actually lead to potential disadvantage of those students most in need of attention. Inclusive targets are formulated in terms of shifting the whole distribution so that no students are left out.

The CoL plans we accessed each had 4-5 achievement challenges, all of which had multiple targets to be achieved by the end of 2017, even when the CoL was in the early stages of formation. One plan had a total of 18 different strategies for reaching the targets, with a similar number for the others. We do not know of any research that would support the idea that forming new collective responsibilities among a number of schools, using multiple strategies, will achieve this number of targets within this time frame. The conditions of persistence in relation to one area with careful monitoring and adjustment of effort over time are not present (Bryk et al., 2015). Indeed, the CoL leaders we talked to who were most advanced in putting the plan into action have made the pragmatic decision to focus on only one of their challenges. However, they were targeting multiple disconnected strategies for doing so.

Given that CoL are unlikely to achieve their targets within the specified timeframes, the second question about who is accountable for what becomes even more interesting. CoL are a voluntary, not a legal, entity. CoL leaders do not have formal authority over individual school leaders. All formal accountabilities, including those in new CoL roles, rest with individual school Boards of Trustees. In an interview with one of the CoL leaders, it was clear that the new arrangements had led to a commitment to collective responsibility within the CoL leadership group (lead principal, across

⁵ Retrieved from <http://www.education.govt.nz/assets/Documents/Ministry/Investing-in-Educational-Success>

school and within school teacher leads), but tensions were developing with some individual principals who were not part of this central structure. These tensions were evident through the across-school leads experiencing difficulty getting into some schools, individual principals believing they should be determining the role in their schools, and others already threatening to withdraw. There is likely to be an ongoing tension between maintaining good relationships across the CoL (to keep them engaged) and holding each other to account for student outcomes. One of the greatest threats Farrar (2015) identified to collective improvement was prioritising relationships over improvement.

This tension is likely to be aggravated with the different accountability frameworks for this major system change and those currently through the governance line of schools' boards of trustees and the Ministry. While the CoL leaders we spoke to were aware of the need for school-to-school collaboration, their definitions of collaboration and the extent to which they saw these as accountability mechanisms among each other were highly variable. In addition, the fundamental tension between incentives for schools to compete with one another (most clearly evident in the mix of school roll size, principals' salaries and parental choice of school), and schools to be collectively responsible for all students in a geographic area, we predict will continue to be problematic.

System Coherence

In 1989 New Zealand replaced a stultifying and overly bureaucratic system with one that favoured greater autonomy for schools and their communities and subsequently greater choice for parents. Since this time, this extreme has moved (appropriately in our view) more into the middle ground. School autonomy and parental choice are now exercised within more constraints. As the system has evolved over the last 26 years, however, the accountability mechanisms appear to have come more as an afterthought than as a central piece of the design of each of the changes. As a result, multiple accountabilities, some serving little purpose beyond providing an account, have led to complexity and duplication.

The OECD review of New Zealand's assessment and evaluation systems (Nusche et al., 2012) commented specifically that our education system is unusually fragmented, and that system coherence is a major issue. The authors noted that this issue exists within and across different education bodies. The Ministry of Education, New Zealand Qualifications Authority, New Zealand Teachers' Council (as it was at the time) and the Education Review Office all have evaluative functions of one kind or another, but there is limited coherence among them, and in some cases, within them. The OECD emphasised the importance of optimising complementarity while reducing duplication and inconsistencies.

One of the difficulties in achieving system coherence is that there does not appear to be a coherent theory for improvement through the system's accountability structures and processes, nor is there an overarching agency responsible for creating a coherent framework across the whole of the schooling system. The new policy emphasis on education for 0-18 years only increases the complexity with the inclusion of early childhood centres that have very different configurations from the schooling sector and very mixed accountability frameworks and mechanisms.

Possible Re-settings

The above analysis highlights some possibilities for re-setting the accountability architecture and systems. In doing this, we are reminded of Suggett's (2015) and Farrar's (2015) analyses referred to in the introduction that, where accountability is weak, autonomy in managing resources and determining curriculum and assessment can in fact worsen rather than improve performance. Our consideration of the re-settings is grouped into two areas: system coherence and the capabilities required in a professionally-oriented accountability system with a strong developmental emphasis.

System Coherence

An analysis of the New Zealand educational context published by the first author of this thinkpiece (Timperley, 2013), prior to the development of CoL, noted the problem of system coherence. In the conclusion to this analysis, Timperley (2013) wrote:

We have an aspirational system with high expectations of the professionals and their communities to achieve these aspirations. For example, teachers are asked to make complex overall judgements about student achievement. Principals are asked to improve teacher performance through navigating between two sets of appraisal standards to attest to their competence. They are also expected to have high quality self-review processes presented in different formats for different agencies. Although accountabilities for performance are weak, each school has multiple accountabilities in the form of reports to a range of agencies and to their communities. All this activity begs the question of whether it is more in the form of busy work or in the form of the kinds of deep change needed to address the real priorities? There is some evidence that busy work trumps deep development. (p.35)

While recent initiatives have attempted to rationalise some of this complexity (e.g. a single set of professional standards), it is clear that this has not yet had sufficient impact.

Each of the accountability systems and mechanisms has been designed with careful attention to the international literature, which emphasises professional accountability with strong links to improvement (e.g. Fullan et al., 2015). However, the fragmentation and duplication has acted as an inhibitor to achieving the goals of improving achievement with greater equity because they increase the administrative burden and distract from the intensive efforts needed to address our enduring educational issues.

As a result, a culture of compliance and reporting has developed, rather than one of accountability with a clear line of sight to improving outcomes for learners, particularly the most vulnerable to educational failure. As one of those we spoke to described when developing this paper, *"It is an unwieldy system with different agencies with different responsibilities that do not connect. There is no shortage of reporting information but there is no mechanism for hanging that information together either for the purposes of accountability or improvement."*

The assessment and evaluation report by the OECD (Nusche et al., 2012) identified the need to map the system. From an accountability perspective, this map needs to include the non-negotiables (e.g. accurate student assessment), what is reported to whom for what purpose, and the consequences that might follow. Rather than viewing the development of internal and external accountabilities as

sequential as suggested by Fullan et al. (2015), they need to be viewed as complementary sides of the coin and brought together into a coherent theory for improvement across the education system.

New Developments and System Coherence

Alongside attempts to rationalise the complexity of reporting and accountabilities, more consideration is needed around how new initiatives interact with existing structures when they are being planned. The development of CoL is a case in point of a major change in the system architecture with limited articulation of how accountability mechanisms for this initiative interact with existing accountabilities beyond reporting. At a superficial level, CoL design is consistent with current international advocacy that improvement follows from collective responsibility and culture change through developing collaboration across groups of schools. But the detail indicates some worrying signs. The short exercise we undertook in examining three CoL plans found some evidence that the collective culture has focused on identifying student deficits and bypasses the contribution of the adults in the system. Clearly, this analysis is insufficient to make substantiated claims and further work is needed to establish the prevalence of these problems. If it is more widespread, then the weight of the research evidence is against the development of CoL leading to improved outcomes for those learners who need to benefit from this system change unless the work in the CoL shifts this emphasis to focus on the contribution of the adults in the system. In addition, it appears that CoL have an unrealistic number of targets to be achieved within an unrealistic timeframe. These issues are likely to detract from rather than promote the serious and difficult work of “raising the average level of quality and performance while at the same time decreasing the variation among units, and engaging people in analysis of and understanding of why some actions seem to work and others don’t” (Elmore, 2004, p. 57).

The interaction between accountabilities in this new entity with existing formal accountability structures remaining with individual schools’ Boards of Trustees is also problematic. To whom are the CoL leaders, across-school and within-school teachers accountable? To whom are the principals of the participating schools who are not in these roles accountable? What are the contingencies when CoL fail to reach their mostly unrealistic number of targets with the timeframe? While logically a possible resetting may involve a stronger role for a CoL governing board in the future, this would serve to add yet another layer of accountability to a system that already requires too many reports to too many people at too many levels unless this governing board becomes a substitute rather than an addition.

It could be argued that the evolving frameworks within which CoL have developed demonstrate a bottom-up or reciprocal accountability as the centre has responded to emerging challenges, including the accountability frameworks. While it is important to learn as implementation progresses, it is important that accountabilities are built into the design as Schleicher (2016) suggests:

There are some basic features of education reforms that students, parents and educators should be able to take for granted. These include clarity of purpose and intended outcomes of reforms at every level of the system. They also include clarity about methods and delivery. Reforms should also include built-in means of public accountability – transparency in when and how judgements will be made as to whether implementation is on track and what the contingency plans are when the intended results do not materialise. Last but not

least, there needs to be clarity about the actors involved in implementation and their relationships. In other words, starting from the policy and ending with the changes in frontline behaviours and practices that this policy is designed to achieve, how – and through whom – does reform actually happen? (pp. 92-93)

We would add that the accountabilities within the reform need to be coherent with other system accountabilities so they enhance their mutual effectiveness rather than detract from them.

Evaluative Capability

New Zealand's strong emphasis on professional accountabilities and linking accountability to development requires high levels of evaluative capability across the system. Evaluative capability means more than setting a range of targets at a system, CoL or school level. It also requires a commitment to testing possible reasons for the current situation focused on those factors over which the key players have control. It answers the question posed in the Treasury document⁶ of "Why is it so?" If the evaluation process is to result in improvement, then those involved need to have the capability to search for research-informed theories for improvement that are designed to address those causal factors. Efforts to make a difference are constantly tested for their efficacy in moving towards agreed goals. It requires collective interpretation and redesign of action in response to the evidence collected (Earl & Timperley, 2015).

We have identified multiple instances throughout this paper where evaluative capability appears to be weak. At times, this problem relates to commitment, for example, some of the politically-motivated reasons given for the assignment of inaccurate overall teacher judgements (OTJs) on National Standards. It is also an issue of culture, with the focus evident in CoL plans being on the deficits of students, rather than considering the adequacy of educational provision. Just as frequently, it appears to be an issue of capability at multiple levels of the system. ERO has identified the problem of 'data literacy' on several occasions.

The important point is that the capabilities required at all levels of the system to maximise success form part of the analysis. Systematic support is then provided to develop the capabilities required. It is often assumed that these capabilities apply to the CoL/schools and kura, but often they are needed at the level of system agencies. Articulating and developing such capabilities must be a priority if a re-conceptualised framework of purposes, mechanisms and processes is to achieve the system shifts needed for better outcomes for Māori and other learners not well served by our education system.

Internationally, much attention is being given to the idea of 'leading from the middle' (OECD, 2015b) to reduce compliance on bureaucracy and to address the important issue of trust required for accountability mechanisms to lead to system improvement. Unquestionably, trust is central, or distortions and avoidance take over from accuracy, energy and motivation. However, trust with limited evaluative capability and clear accountabilities can often lead to maintenance of the status quo, particularly in systems with high levels of autonomy (Farrar, 2015). It is easier to do what we have always done, and maybe tweak a few things around the edges while maintaining collegial relationships, than to fundamentally rethink how we go about creating a learner-focused system in which all with responsibility for educational provision are accountable for the outcomes, particularly

⁶ The Treasury (undated). *Higher Living Standards and Education*

those currently least well-served by the system. Much more is required to create the system change needed to improve outcomes for our most vulnerable learners. The development of many more capabilities are required.

Where next?

A system map of existing accountabilities would help to identify the duplications and multiple reports identified in this analysis. This map would also need to identify the purposes and capabilities required in order for the accountabilities to make a difference to priority learners. Redesign, however, would require much more. To begin, a shared definition of accountability would help: is it synonymous with responsibility (Fullan et al., 2005) or the remainder when responsibility is subtracted (Hargreaves, 2008)?

If accountability mechanisms and processes are to benefit those learners who are not well-served by our education system, the redesign could start from a student perspective and work up through the system, rather than starting with national agencies with each deciding what is good for others. Some questions might include:

- How can reporting be positioned to promote rather than detract from improvement in outcomes for our most vulnerable learners?
- How can collaborative cultures that emphasise professional accountability be developed throughout the system and what capabilities are needed where for them to be effective?
- How can the question, 'Why is it so' be asked and answered at all levels of the system in ways that identify what needs to change?
- What kinds of expectations and consequences will follow from the different accounts?
- How can the accountabilities be inclusive of a system from 0-18 years, not separately for each sector?

References

- Bendikson, L. (2015). *Community of Schools' (CoSs) leadership: Throwing money and hoping for success?* Think piece for Education Council of New Zealand. Retrieved from <https://educationcouncil.org.nz/sites/default/files/Five%20Think%20pieces.pdf>.
- Bishop, R., Berryman, M., Cavanagh, T., & Teddy, L. (2009) Te Kotahitanga: Addressing educational disparities facing Māori students in New Zealand. *Teaching and Teacher Education* 25(5), 734–742.
- Blank, A., Houkamau, C., & Kingi, H. (2016). *Unconscious bias and education: A comparative study of Māori and African American students*. Auckland: Oranui.
- Bryk, A., Gomez, L., Grunow, A. & LeMahieu, P. (2015). *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Cambridge, Massachusetts: Harvard Graduate School of Education
- Bryk, A.S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Sage.
- Daly, A.J., & Chrispeels, J. (2008). A question of trust: Predictive conditions for adaptive and technical leadership in educational contexts. *Leadership and Policy in Schools*, 7(1), 30–63.
- Darling-Hammond, L. (2014). One piece of the whole: Teacher evaluation as a part of a comprehensive system for teaching and learning. *American Educator*, 38(1), 4-13.
- Earl, L., & Timperley, H. (2015). *Evaluative thinking for successful educational innovation*. OECD Education Working Papers, No. 122. Paris: OECD Publishing.
- Education Council (2016). *Strategic options for developing future orientated initial teacher education*. Wellington: Education Council. Retrieved from <https://educationcouncil.org.nz/sites/default/files/Strategic%20Options%20for%20Developing%20Future%20Orientated%20Initial%20Teacher%20Education.pdf>.
- Education Review Office (2016). *School evaluation indicators: Effective practice for improvement and learner success*. Wellington: Education Review Office.
- Education Review Office (2014). *Raising achievement in primary schools*. Wellington: Education Review Office. Retrieved from <http://www.ero.govt.nz/publications/raising-achievement-in-primary-schools/>.
- Ell, F., Mayo, S., Twyford, K., Meissel, K., Yao, E., McIlraith, G., Le Fevre, D., Si'ilata, R., & Timperley, H. (2015). *The influence of standardised test results on overall teacher judgments in writing*. CPL report. Commissioning body: Consortium for Professional Development.
- Elmore, R. (2003). *Knowing the right thing to do: School improvement and performance-based accountability*. Washington, DC: NGA Center for Best Practices.
- Elmore, R. (2004). *School reform from the inside out: Policy, practice and performance*. Cambridge, MA: Harvard University Press.
- Elmore, R. (2005). Accountable leadership. *The Educational Forum*, 69(2), 134-142.

- Farrar, M. (2015). *Learning together: The power of cluster-based school improvement*. Melbourne: CSE.
- Fullan, M., Rincon-Gallardo, S., & Hargreaves, A. (2015). Professional capital as accountability. *Education Policy Analysis Archives*, 23(15), 1-22.
- Greany, T. (2015). More fragmented, and yet more networked: Analysing the responses of two Local Authorities in England to the Coalition's 'self-improving school-led system' reforms. *London Review of Education*, 13(2), 125-143.
- Hargreaves, A. (2008). The fourth way of change: Towards an age of inspiration and sustainability. In A. Hargreaves & M. Fullan (Eds.), *Change wars* (pp. 11–44). Toronto: Solution Tree.
- Hargreaves, D. H. (2012). *A self-improving school system: Towards maturity*. Nottingham: National College for School Leadership.
- Hattie, J. (2005). *What is the nature of evidence that makes a difference to learning?* ACER Conference Paper. Retrieved from http://research.acer.edu.au/research_conference_2005/7.
- Hattie, J. (2015). *What works best in education: The politics of collaborative expertise*. London: Pearson.
- Hynds, A., Averill, R., Hindle, R., Meyer, L. (in press). School expectations and student aspirations: The influence of schools and teachers on indigenous secondary students. *Ethnicities*.
- Katz, S. Earl, L., Ben Jaafar, S., Elgie, S., Foster, L., Halbert, J. & Kaser, L. (2008). Learning networks of schools: The key enablers of successful knowledge communities. *McGill Journal of Education*, 43(1).
- Kiro, C., Hynds, A., Eaton, J., Irving, E., Wilson, A., Bendikson, L., Cockle, V., Broadwith, M., Linley-Richardson, T., & Rangitiki, M. (2016). *Starpath Phase 2: Final evaluation report*. Auckland: Starpath Project, the University of Auckland.
- Kogan, M. (1986). *Education accountability: An analytic overview*. London: Hutchinson.
- Lasky, S., Schaffer, G., & Hopkins, T. (2009). Learning to think and talk from evidence: Developing system-wide capacity for learning conversations. In L. Earl & H. Timperley (Eds), *Professional learning conversations: Challenges in using evidence for improvement*, (pp. 95-107). Dordrecht: Springer.
- Meissel, K., Meyer, F., Yao, E., & Rubie-Davies, C. (in review). Subjectivity of teacher judgments: Exploring student characteristics that influence teacher judgments of student ability. *Teaching and Teacher Education*.
- Ministry of Education (2010). *OECD review on evaluation and assessment frameworks for improving school outcomes*. Wellington: Ministry of Education.
- Ministry of Education (2011a). *Ministry of Education position paper: Assessment [schooling sector]*. Wellington: Ministry of Education.

- Ministry of Education (2011b). *Tātaiako: Cultural competencies for teachers of Māori learners*. Wellington: Ministry of Education.
- Ministry of Education (2013). *The Māori education strategy: Ka Hikitea – Accelerating success 2013-2017*. Wellington: Ministry of Education.
- Ministry of Education (2015a). *Community of schools tips and starters: Working together*. Wellington: New Zealand Government.
- Ministry of Education (2015b). *The New Zealand curriculum*. Wellington: Learning Media.
- Ministry of Education (2016). *Community of learners: Guide for schools and kura*. Wellington: Ministry of Education. Retrieved from <http://www.education.govt.nz/assets/Documents/Ministry/Investing-in-Educational-Success/Communities-of-Schools/Communities-of-Learning-Guide-for-Schools-and-Kura-web-enabled.pdf>.
- Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved systems keep getting better*. New York: McKinsey & Company.
- Munby, S., & Fullan, M. (2016). *Inside-out and downside-up: How leading from the middle has the power to transform education systems*. Think/action piece for the Education Development Trust and Motion Leadership.
- Mutch, C. (2013). Developing a conceptual framework for school review. In M. Lai & S. Kushner (Eds.) *A Developmental and Negotiated Approach to School Self-Evaluation* (pp. 91-108). Bingley: Emerald Group Publishing.
- Nusche, D., Laveault, D., MacBeath, J., & Santaigo, P. (2012). *OECD reviews of evaluation and assessment in education: New Zealand 2011*. Paris: OECD Publishing.
- OECD (2010). *PISA 2009 results: What makes a school successful? Resources, policies and practices (Volume IV)*. Paris: OECD Publishing.
- OECD (2011). *Improving school leadership policy and practice: Pointers for policy development*. Paris: OECD Publishing.
- OECD (2012). *Education today 2013: The OECD perspective*. Paris: OECD Publishing.
- OECD (2013a). *PISA 2012 results: What makes a school successful? Resources, policies and practices (Volume IV)*. Paris: OECD Publishing.
- OECD (2013b). *Synergies for better learning: An international perspective on evaluation and assessment*. Paris: OECD Publishing.
- OECD (2014). *Improving schools in Wales: An OECD perspective*. Paris: OECD Publishing.
- OECD (2015a). *Improving schools in Scotland: An OECD perspective*. Paris: OECD Publishing.

- OECD (2015b). *Schooling redesigned: Towards innovative learning systems: Educational research and innovation*. Paris: OECD Publishing.
- OECD (2016). *PISA 2015 assessment and analytical framework: Science, reading, mathematics and financial literacy*. Paris: OECD Publishing.
- Phillips, G., McNaughton, S., & MacDonald, S. (2001). *Picking up the pace: Effective literacy interventions for accelerated progress over the transition into decile 1 schools*. Final Report to the Ministry of Education on the Professional Development associated with the Early Childhood Primary Links via Literacy (ECPL) Project. Auckland, New Zealand: The Child Literacy Foundation and the Woolf Fisher Research Centre.
- Robinson, V., & Timperley, H. (2000). The link between accountability and improvement: The case of reporting to parents. *Peabody Journal of Education*, 75(4), 66-89.
- Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change*, 11(1), 45-61.
- Schleicher, A. (2016). *Teaching excellence through professional learning and policy reform: Lessons from around the world*. Paris: OECD Publishing.
- Schütz, G., West, M., & Wößmann, L. (2007). *School accountability, autonomy, choice and the equity of student achievement: International evidence from PISA 2003*. OECD Education Working Paper. Paris: OECD Publishing.
- Sinnema, C. (2014). *Evaluation indicators for school reviews: A commentary on the effective teaching indicators*. Report prepared for Education Review Office Think Tank.
- Suggett, D. (2015). School autonomy: Necessary but not sufficient. *Evidence Base*, 1, 1-33.
- Tetlock, P.E. (1998). Losing our religion: On the precariousness of precise normative standards in complex accountability systems. In R. M. Kramer & M. A. Neale (Eds.), *Power and influence in organizations* (pp. 121-144). Thousand Oaks: Sage.
- Timperley, H.S. (2013). The New Zealand educational context: Evaluation and self-review in a self-managing system. In M. Lai & S. Kushner (Eds). *A developmental and negotiated approach to school self-evaluation* (pp.95-108). Bingley: Emerald Group Publishing.
- Timperley, H. (2014). *Promoting and supporting improvement in schools through external review*. (Paper prepared for the Education Review Office). Wellington: Education Review Office.
- Timperley, H., Hohepa, M., Keegan, P., Parr, J., Lai, M., & McNaughton, S. (2010). *Building Evaluative Capability in Schooling Improvement: Milestone Report, Part A: Strands One and Strand Three*. Report to the Ministry of Education, Wellington, New Zealand.
- Timperley, H., Kaser, L., & Halbert, J. (2014). *A framework for transforming learning in schools: Innovation and the spiral of inquiry*. Melbourne: CSE.
- Timperley, H., & Parr, J. (2009). Chain of influence from policy to practice in the New Zealand Literacy Strategy. *Research Papers in Education*, 24(2), 135-154.

- Timperley, H.S., Parr, J.M., & Bertanees, C. (2009). Promoting professional inquiry for improved outcomes for students in New Zealand. *Professional Development in Education*, 35(2), 227-245.
- Timperley, H.S., & Robinson, V.M.J. (1998). Collegiality in schools: Its nature and implications for problem solving. *Educational Administration Quarterly*, 34, 608-629.
- Timperley, H.S., & Wiseman, J. (2003). *In-school processes related to the sustainability of professional development in literacy*. Report to the Ministry of Education, Research Division.
- Tucker, M. (2014). *Fixing our national accountability system*. Washington D.C: The National Center on Education and the Economy.
- Turner, H., Rubie-Davies, C. M., & Webber, M. (2015). Teacher expectations, ethnicity and the achievement gap. *New Zealand Journal of Educational Studies*, 50(1), 55-69.
- Ward, J., & Thomas, G. (2016). *National Standards: School sample monitoring and evaluation project, 2010-2014*. Report to the Ministry of Education. Wellington: Ministry of Education.
- Wößmann, L., Lüdemann, E., Schütz, G., & West, M. (2007). *School accountability, autonomy, choice, and the level of student achievement: International evidence from PISA 2003*. OECD Education Working Paper, No. 13. Paris: OECD Publishing.
- Wylie, C. (2012). *Vital connections: Why we need more than self-managing schools*. Wellington: New Zealand Council for Education Research.
- Wylie, C., & Bonne, L. (2014). *Primary and intermediate schools in 2013: Main findings from the NZCER national survey*. Wellington: NZCER.

A systems perspective on accountabilities in the school system

Draft 1

Graham Scott

31 August 2016

Introduction

This report responds to the terms of reference provided by the Treasury to:

Provide a short think piece on school accountability, which provides an answer to the following questions:

- What is your perspective on how the accountability settings in New Zealand's schooling system should be set to encourage and enable teachers, principals and schools to work in ways which drive sustained improvement in student's learning outcomes? In what ways can accountability settings specifically help to improve learning outcomes for those children most at-risk of not achieving?
- Think about the purposes of the accountability system, including what behaviours the accountability system is aiming to motivate and enable among principals and teachers.
- What do you think are the key accountability levers at different levels of the system, and what is your assessment of what should change?
- Think about the appropriate mix of 'hard' and 'soft' accountability levers. What is the government's role in building capability and creating the conditions in which professional autonomy and accountability drives continual system improvement.
- This work is in the context of the importance that the Treasury places on Māori outcomes and on accelerating Māori educational achievement.

The assignment asked for me to respond from the perspectives of my experience in public management and more specifically my work in designing the Integrated Performance and Improvement Framework, which with some modifications and simplifications has been absorbed into the Government's recently refreshed health strategy. This was one example of an attempt to design an accountability and performance framework for a complex system. The school system is less complex than the health system but there are insights from systems thinking about health that are relevant. This report does not attempt a theoretical discussion of complex adaptive systems but rather some practical insights that point to sources of inefficiency in present arrangements that might be remedied by measures suggested from that perspective.

The report begins with a note on the current policy context in which this work in Treasury is being conducted. A short conceptual discussion about accountabilities and their place within a strategy for system wide performance improvement follows and includes a brief description of the Integrated Performance and Improvement Framework in the health system. After a discussion of the current performance of the school system the report reflects on the objectives for the system and how they are changing. The report returns to comment on the specific questions in the TOR with some final conclusions.

Context

Perspectives on the performance of the school system display both light and shade. Material of a general nature on the Ministry's website is very positive, but with the admission that "alongside these points of strength, New Zealand is challenged by a significant tale of low achievement among some groups which we continue to strive to address"¹. OECD and other data showing high rankings in general but slipping as other countries move past us stop there is continuing unsatisfactory performance in lower socio-economic groups and among Maori and Pacifica. There are concerns about the current performance of the public school system - including the decile system - and about the fitness of the system to succeed within a changing national and global context.

More specific performance data produced by the Ministry of education is generally very upbeat about achievements². On the other hand, the report of the Task Force on Regulations Affecting School Performance was very critical of the system. It considered that the Education Act does not currently provide a clear and visible purpose statement that articulates the broad goals and outcomes desired of the New Zealand schooling system³. While designed to provide such high level direction, the National Education Goals (NEGs) were and are not currently achieving this purpose in the view of the Task Force and others.

The report says the current planning and reporting process is not fit for purpose and has not kept up with the evolution of school management.

Along with many other criticisms in the report, these two criticisms above in particular amount to saying that the goals of the system are wrong and not achieving their purpose, while the planning and reporting processes, which are the skeleton of the accountability system, are failing.

The Education Act by these accounts is obsolete and so this is an important time to learn the lessons of decades of experience since the Picot reforms and with the structures and processes that were put in place to support it. This review by Treasury of the system of accountability is an opportunity to re-think the concepts and use of accountability mechanisms in the interests of improved system-wide performance. Presumably this will feed into Treasury's advice on the reform of the education Act currently underway.

Accountabilities as an element of organisational architecture

The TOR takes a fairly broad view of accountabilities noting that they may be hard or soft and collectively constitute a system. There is a presumption that some set of accountabilities can be devised that drive sustained improvement by the school system as a whole. Perhaps there is, but it for any accountability system to be both necessary and sufficient to drive system-wide performance improvement would require the other important influences on performance to be aligned with the objectives of the accountabilities. This is unlikely to be the case and the TOR acknowledge this in relation to one important influence on performance, specifically that capability and 'conditions' may not be favourable to performance improvement through the exercise of autonomy and accountability.

¹ <http://www.education.govt.nz/assets/Uploads/NZ-Education-System-Overview-publication-web-format.pdf>

² <http://www.education.govt.nz/news/information-is-key-to-improving-education-performance/>

³ Ten National education goals <http://www.education.govt.nz/ministry-of-education/legislation/negs/>

This invites analysis of what is a broadly-defined accountability system for school performance and what are the circumstances – if any - in which it is likely to drive sustained improvement in the performance of the state school system. What complementary capabilities, values, motivations, resources etc are likely to be needed in order for an accountability system to drive sustained performance? Or to foreshadow the perspective in this report, how does an accountability system align with the other factors that together promote sustained performance improvement?

Also there is the question of the scope of these accountabilities. Are they predominantly concerned with the accountability of the recipients of delegated resources and the authorities from the State for the exercise of those delegations and resources? Or does the system of accountabilities include those between agents who are not directly accountable to the State or at least have multiple accountabilities to the State and to other parties whose interests may differ from the State's. In so far as the TOR raise the question of professional accountability it would seem so.

The fundamental meaning of accountability is that a person or organisation is expected or required to account for themselves in respect to their conduct of some duty. Being accountable is essentially to explain oneself and possibly, but not necessarily, face consequences either pleasant or unpleasant. The common call for someone to be held accountable is a call for someone to be punished, but this is just one dimension of an accountability relationship and not universal.

People typically think of an accountability as a relationship based on asymmetric power between two parties. The person with the power is holding the other to account. But accountability relationships can also be mutual, as for example in 'collective impact' models of service delivery. In these, two parties to a joint venture or similar arrangement are accountable to each other to make each's agreed contribution to the shared endeavour. Such arrangements also extend to multi-party teams in pursuit of common goals as for example in the better functioning 'alliances' in the health sector. Where parties come together to pursue shared goals that are yet to be defined clearly or may never be, then the accountability can take the form of commitment to values and processes rather than specified goals. This can be found in the open source movements in the tech and related industries. A crucial element of the success of these kinds of accountability relationships is trust.

The question of trust is crucial to the conduct of accountability relationships especially where multiple parties are involved in an endeavour, where attribution of results to individual participants in team is difficult and where goals are not singular or are difficult to pin down and to measure. In these circumstances there is great scope for free riding, shirking and opportunism in the conduct of the accountability relationship. In a devolved system and one in which the 'product' is a post-experience or credence good, the skills of seasoned professionals on the front line to make decisions that are tailored to unique individuals or groups drive the outcomes. The accountability system has to give high levels of trust and discretion to them while also ensuring that incentives arising from the conditions under which they work do not distort their behaviours away from the outcomes that are sought.

Even in the one-on-one model in classic agency theory there are limits to the effectiveness of contracts between principals and agents in aligning the interests of the agent with those of the principal. All contracts are incomplete in the sense that not all possible future contingencies can be allowed for. Hence there are residual risks not anticipated by the principal-agent agreement that fall on one party or another. The attempt to tie down contractually all the possibilities accumulates 'agency costs' up to a point where alternatives to contractual detail are superior for constraining opportunism by the agent. Typically the alternative is employment of the agent by the principal, with the implication that the principal can direct the agent as events unfold.

These agency costs are magnified where there are multiple agents that must work together or where agents have additional accountabilities to other parties. Costs are magnified again when the objectives that the principal seeks to motivate the agent to achieve are multiple or vague. So, in a multi-faceted system like the State school system, it is intuitively apparent that recourse to a simple principal-agent concept of accountability is going to be difficult to fine tune in the pursuit of system-wide performance goals.

Accountabilities in complex adaptive systems

The simple agent-to-principal accountability concept – while useful for many purposes – is not well-suited to situations where:

- complex systems and networks are involved in the scope of enquiry
- where the desired objective or outcome is multi-dimensional and not able to be summarised in simple targets or metrics
- where the provider of the service at the point of delivery is making fine professional judgements on complex and individualised information about the client and circumstances
- where the 'recipient' of the service is engaged in co-production with the service deliverer to achieve joint objectives
- where the outcome requires collaboration of numerous parties to reach the result
- where knowledge of what works in achieving the outcome is contingent, circumstantial and changing

Unfortunately for the designers of accountability regimes these conditions apply where the system that is to be controlled is a complex adaptive system.

Complex adaptive systems can be defined in terms of the following characteristics⁴:

- "They are *nonlinear and dynamic* and do not inherently reach fixed-equilibrium points. As a result, system behaviours may appear to be random or chaotic.
- They are composed of *independent agents* whose behaviour is based on physical, psychological, or social rules rather than the demands of system dynamics.
- Because agents' needs or desires, reflected in their rules, are not homogeneous, their *goals and behaviours are likely to conflict*. In response to these conflicts or competitions, agents tend to adapt to each other's behaviours.
- Agents are *intelligent*. As they experiment and gain experience, agents learn and change their behaviours accordingly. Thus overall system behaviour inherently changes over time.
- Adaptation and learning tend to result in *self-organization*. Behaviour patterns emerge rather than being designed into the system. The nature of emergent behaviours may range from valuable innovations to unfortunate accidents.
- There is *no single point(s) of control*. System behaviours are often unpredictable and uncontrollable, and no one is "in charge." Consequently, the behaviours of complex adaptive systems can usually be more easily influenced than controlled.

In these systems not all system design and management problems can be addressed through hierarchical decomposition. For example, decomposition may result in the loss of important information about interactions among the phenomena of interest. Another fundamental problem for very complex systems like health care is that no one is "in charge," no one has the authority or

⁴ Rouse, W.B., Health Care as a Complex Adaptive System: Implications for Design and Management, The Bridge, National Academy of Engineering, Spring 2008

resources to design the system. Complex adaptive systems tend to have these design and management limitations.”

It is widely accepted that health systems display these characteristics and that efforts at performance improvement should take them into consideration. The Integrated Performance and Improvement Framework developed in 2013⁵ attempted to do so. Annex 1 contains two slides that capture the essence of the design. The Government sets high level goals for the health sector drawing on the ‘triple aim’ and supported by health goals for the main demographics. The providers are brought together in shared goal models of collaboration in ‘Alliances’, which are brought together by DHBs but not ruled by them. The Canterbury Clinical Alliance is generally regarded as the most successful of these so far. Agreements are reached within the alliances as to the most appropriate ‘pathways’ through the system for people presenting with various symptoms. Clinical leaders, not health bureaucrats, govern and control the pathways. All the participants to each pathway agree mutually on what the pathway should be and to play their part in implementing it.

There is a second governance structure under the minister and the Director General of Health to provide for the necessary control and accountability for the implementation of government health policy and accounting for the use of resources and other statutory requirements.

In order to forge a link between national health goals and the business models within the alliances a lot of work was done by clinicians to develop a dictionary of performance indicators from which they could draw whichever ones they thought most appropriate to plan and monitor their contribution to the broader national goals. This solves the problem of ad hoc indicators using different definitions and different data sources, which undermines comparability and the diffusion of good practices. Alliances also have some freedom to use other indicators where in their judgement these were more appropriate to the circumstances of individual regions – for example higher Maori populations in some regions.

The pressure for improved performance can be seen in the rows across the diagram in the annex. The simple idea is that all participants should pass entry level standards or be excluded, upper rows capture higher levels of performance by alliances and the top level provides for investments in innovation of new modes of service provision.

The bones of the proposal have survived into the revised health strategy announced by the Government. It is simplified and the performance incentives diminished. But the feature of using the dictionary of subsidiary performance indicators is attracting growing support.

The question for this report is whether the New Zealand school system is also some kind of complex adaptive system. It has many of these characteristics but also intuitively seems simpler than the health industry. If so, then it has important implications for what should be in an accountability regime to drive up system performance and whether this is a realistic prospect. If not then what instead or in addition is needed to improve performance?

Key features of this design for better integrated health and social services and that are probably relevant to the state education system are:

- Governance and control of the accumulating body of clinical practice is in the hands of professional experts

⁵ <http://www.hiirc.org.nz/page/45527/>

- High level system goals are fairly general but sufficiently clear as to motivate service providers to arrange their businesses so as to make measurable contributions to the higher level goals through the achievement of performance indicators that they choose from a dictionary of generally accepted indicators
- There are arrangements for collective impact by an alliance of providers to achieve jointly agreed goals using agreed pathways and clinical evidence
- Accountability is administered to the collectives through leadership roles that may or may not be state agencies or key personnel.

System architecture

Judgements about particular accountabilities need to be grounded in knowledge of the institutional arrangements that apply to the whole system being studied. This is essential to avoid using accountabilities in the way described in the following quote:

“A third approach (to attaining transformation in education), usually advocated by politicians, is to make systems more strongly accountable for learner performance, guided by a belief that somehow someone will know how to do this well and will make the accountability – this time – really count⁶.”

Whether a particular accountability mechanism makes a positive or negative contribution depends largely on its alignment or misalignment not only with other accountabilities within the same system but with other institutional features that affect the governance and performance of the system. For want of a better name I am calling these institutional arrangements the “system architecture”. This is an extrapolation of the concept of organisational architecture applying to a single organisation in order to take account of the institutional arrangements for a system involving numerous organisations.

Organisational architecture is a metaphor drawn from conventional architecture to provide a framework through which an organisation expresses its vision, strategy, culture and values while providing for its infrastructure, formal organisational structure, governance, roles and delegations, business processes and human resources.

Extrapolating the concept from an organisation to a system necessarily sacrifices some clarity, but serves to reinforce the important point that the meta-architecture of the system incorporates by way of mutual interaction the architecture of the constituent organisations within the system. As the remarks about complex adaptive systems above highlight, the performance of the system as a whole is not just a summation of the performance of the individual parts. The interaction between the parts creates dynamics and feedbacks which can have a major effect on whole-of-system performance.

Accountabilities in the school system

A scan of the present system of accountabilities reveals features that seem unsatisfactory from the conceptual perspective on accountability systems above.

⁶ Helen Timperley, Linda Kaser, and Judy Halbert, A framework for transforming learning in schools: innovation and the spiral of enquiry, Centre for Strategic Education, April 2014, p3.

Goals

The goals for the school system are numerous and not apparently consistent or aligned. The National Educational Goals in the act are, as noted above, heavily criticised by the Task force on Regulations.

“the Act does not currently provide a clear and visible purpose statement that articulates the broad goals and outcomes desired of the New Zealand schooling system. While designed to provide such high level direction, the National Education Goals (NEGs) are not currently achieving this purpose.”

The Taskforce recommends that: “the Act contain a purpose statement outlining the desired outcomes for the schooling system this statement be enduring, inclusive, student-centred and embrace a breadth of desired student outcomes” But says it should be developed with the sector.

The bill to update the Education Act introduced this month proposes new goals as follows⁷:

- helping each child and young person attain educational achievement to the best of their potential; and
- promoting the development of:
 - resilience, determination, confidence, and creative and critical thinking
 - good social skills and the ability to form good relationships
 - participation in community life and fulfilment of civic and social responsibilities
 - preparedness for work
- to instill an appreciation of the importance of:
 - the inclusion within society of different groups and persons with different personal characteristics
 - the diversity of society
 - cultural knowledge, identity, and the different official languages
 - the Treaty of Waitangi and te reo Maori

Whilst these are very general and rather vague, they are intended to be timeless goals that endure as different governments impose their own specific priorities on the system (emphasis added).

“Minister of Education (the Minister) will be authorised to issue a statement of National Education and Learning Priorities which **school boards of trustees will be required to take into account** in their teaching and learning programmes and be accountable through their planning and reporting.”

“**Boards will be required to give effect** to the National Education and Learning Priorities when working towards their objective of ensuring that each student is able to achieve the highest standard of education and learning.

Sponsors of Partnership Schools will be required to ensure that, in the development and delivery of any curriculum by a school, the school **has regard to** the National Education and Learning Priorities.

⁷ <http://www.education.govt.nz/ministry-of-education/legislation/the-education-update-amendment-bill/establishing-enduring-objectives-for-children-and-young-people-in-education/>

Communities of Learning that have an agreement with the Secretary for Education will be required to **take it into account** the National Education and Learning Priorities when developing a plan.”

The highlighted words demonstrate the difficulty of expressing the accountability of the named ‘agents’, who all have substantial autonomy, for delivering the various requirements. To complicate matters further, the National Administrative Guidelines and various other guidelines imply another set of objectives and requirements. The Better Public Service Goals for education are a layer over the top of the system, not necessarily consistent with the statutory goals and other guidelines and possibly creating unintended negative side effects.

It should be no surprise therefore that the central authorities are challenged in getting a system of accountabilities that they believe will drive up sustained performance improvement towards specified goals. However the rather vague ways in which the accountabilities above are expressed for delivering on ‘priorities’ – undefined – that ministers will set from time to time, reflects the reality that each of the entities listed above have various and high degrees of autonomy. In other areas of public policy and legislation phrases like “give effect to” have been defined by the courts in landmark court cases, e.g. the Supreme Court Judgement on the King Salmon case⁸. The vague accountabilities in the school system – if they are to be most effective – will require definition and processes to support them in practice.

It is a reasonable conclusion that it will remain difficult to get strong leverage on system performance through an accountability system that holds semi-autonomous bodies to account for very general educational goals, overlaid with more specific and changing priorities at several levels of authority. The question of ‘accountability for what’ needs a more precise answer.

Current accountability mechanisms

The Treasury background documents describe the accountabilities and current proposals to improve them as follows:

1. From Ministry of Education to Government

Ministry of Education (MoE) provides advice and reports directly to the Minister of Education.

Treasury supports Ministers to hold central Education agencies to account through planning and strategic processes; budget and financial management processes and overview of reform processes.

The State Services Commission supports Ministers to hold central Education agencies to account through: reporting against BPS targets; CE performance management and Performance Improvement Framework (PIF) reviews.

2. From schools to central education agencies.

School charters are reviewed by MoE to ensure compliance with the Act. Schools are required to provide audited Annual Reports, National Standards and other data to the Ministry. The Ministry has an intervention framework for significant under-performance.

School evaluation integrates schools internal evaluation and ERO’s external evaluation. The Education Council sets registration and certification standards and processes; new Code of Conduct; and disciplinary processes.

⁸ <https://www.courtsofnz.govt.nz/cases/environmental-defence-society...>

The Act update will establish a strategic framework for education and streamline the planning and reporting process. It will improve transparency around a set of key measures and strengthen the Ministry's ability to support schools that are struggling.

3. From schools to parents

Parents elect the Board of Trustees, which has school governance responsibilities.

Parents receive information about children's achievement (NS, NCEA) and access to school information (ERO reviews, school charter and annual report). This will be strengthened through additional transparency and reporting.

4. From professionals to professionals

Communities of learning aim to motivate and enable continued improvements in teaching and school effectiveness through collaboration.

The Education Council has a new professional leadership function.

Comment

With the exception of item 4, which is too vague to be cast as an accountability, these accountabilities appear to be grounded in a hierarchal conception of the school system based on planning, budgeting, reporting, auditing and intervening to correct departures from expectations. There is nothing in this list to suggest that they emerge from a system wide consensus on what would put the school system on a path of sustained improvement and attention to issues of equity and access. This is a very government centric view of the school system, which is an impression that is reinforced by the emphasis on government instrumentalities in official documents:

- Ministry of Education
- Careers New Zealand
- the Tertiary Education Commission
- the New Zealand Qualifications Authority
- the Education Review Office
- Education New Zealand, and
- the Education Council (an independent professional body).

It is vital that the centre of the government ensures that its own organisations are functioning well and aligned to system wide education goals. However, the lack of visibility of an underlying agreement about how to achieve better system wide outcomes, and the hierarchal feel of the accountabilities above, are reason to be concerned about the misuse of accountabilities as described above.

Add to this some reasons for concern over some of the incentives which are at work in the system, results in reason to about and how these broad accountabilities play out in and around the classroom. As the Novopay problem illustrated, there are an enormous range of minor allowances in teacher pay. Principals routinely recommend essentially all teaching staff for progression through the collective agreement. These are evidence of a very distorted HR system, which must impact on teacher performance and orientation.

There is a lot of central government administrative machinery that is overseeing 2500 schools. And while schools are generally easier to run than health facilities, it is notable that, by comparison, the school system seems lacking in support for schools to implement what is known to work in raising school system performance. There must be huge variations in capability across that number of

schools in great variety in the challenges they face. The most crucial role of any school board is the appointment of the principal, but it cannot be assumed that this number of boards all have the requisite skills to be able to identify principals who can drive up the performance of the school.

The Education Act has a lot of flexibility within it and gives schools considerable freedom in how they conduct their affairs and even how they teach the national curriculum. But is there a nexus of accountability expectations and capability limitations that has resulted in these freedoms not being as widely used as might be expected?

The schools are feeding up to the government accountability agencies information about their affairs that is interest to the latter. But is there a flow of information about effectiveness in terms of pedagogy, literacy, numeracy, student engagement and safety and engagement with parents and community. If it is thought that these are what makes a difference about performance then you would expect to see more emphasis on these things in the systems of goalsetting, reporting and accountability.

There is evidence that some of what is reported is not seen as being particularly useful either to the schools or to the agents of accountability. This is scarcely surprising, as consultations by the Productivity Commission as part of its report on innovations and social services, showed this to be a pattern across social policy. How valuable for example are the School Charters? What is done with them that makes a difference to what happens in class rooms?

ERO does well in evaluating schools, but the evidence is clear that there is a very wide variation in student achievement around the average. Some students do very well in decile 1 schools while others do poorly in decile 10 schools. The strategies for addressing long tails of underachievement are demonstrably failing. If addressing this is a priority you would expect to see clearly visible within the system information about what results are being achieved with these children and how a body of learning is being accumulated through innovation and imitation.

The summary of the changes to be made under the education amendment bill struck me as generally managerial, by way of clarifying roles and accountabilities, improving planning and reporting, more graduated interventions, communities of learning, cohort entry, compulsory attendance, managing enrolment schemes etc. The material on improving accountabilities states⁹:

Strengthen the efficiency, effectiveness and accountability of schools and the schooling network

Align the strategic direction-setting with strong accountability through defined roles and responsibilities for school boards of trustees, and meaningful planning and reporting. Emphasise the importance of collaboration for a sustainable, high performing education system.

[Clarifying Boards of Trustees' Roles and Responsibilities](#)

[Improving planning and reporting](#)

Put in place additional interventions to enable faster, more tailored responses when a school is struggling to ensure the achievement of all its children and young people.

[Creating a more graduated range of interventions](#)

⁹ <http://www.education.govt.nz/ministry-of-education/legislation/the-education-update-amendment-bill/>

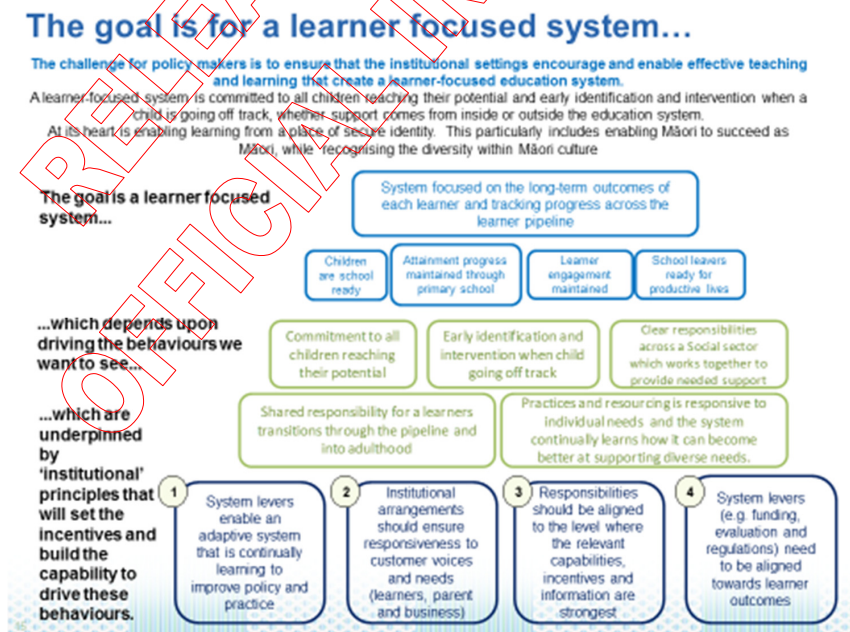
if the impression above is fair, then there is reason to address the question of what the system architecture would look like and how would it function if it were flipped, so that individual student learning is at the centre and the myriad levers through the system are aligned to achieving measurable improvement with that goal. The high variance in achievement across schools and deciles would be targeted for reduction, and in particular to address the pupils in the long tails.

Sustained performance improvement

This report is not intended to assess alternative national education strategies and pedagogical philosophies and methods. However it is useful to provide a caricature of a more desirable system as scheduled the background Treasury documents. This will complete the basis from which to make concluding comments about the questions in the terms of reference.

In these documents, the education system is described more broadly than is typical by extending its coverage to include employers, communities and other social service providers. The state functions are described as system stewards, funders, governors, regulators and information providers. These functions are alongside the other players rather than at the top of the hierarchy, as is normally seen in organograms. The inclusion of other social service providers is a necessary response to the fact that many students in the long tails and their families are in situations where they need other social services. Also when those services are delivered through a silent pipeline they can accidentally disrupt a child's education e.g. by shifting the family so the child has to attend another school.

The goal of the system is to be focused on each learner and the means for doing so are summarised in the following slide:



The row at the bottom of this slide depicts the system architecture which is the focus of this report. To develop the system architecture to support this would require detailing objectives and risks within these elements, business processes to meet these objectives leading onto descriptions of

necessary capability, roles, delegations, reporting requirements, learning and evaluation to make those business processes successful together with governance and accountability relationships.

Business processes would be founded partly on a philosophy of pedagogy in student engagement. An emerging view of the student-teacher relationship takes account of the coproduction in which both make an effort to achieve a common goal¹⁰. This concept is useful in thinking about what a learner focused system might involve. It is helpful in thinking about how to address incentive distortions that may explain some of the long tail e.g. NCEA passes in subjects which do not provide a basis for a tertiary education and student-teacher gaming of various kinds.

The paper quoted above by Timperley et al also has useful insights on how a learner focused system might work.

Accelerating Maori educational achievement

The place of accountabilities within a system that yields improved outcomes for Maori pupils should be derived from the institutional architecture that has been chosen for the purpose. A particular accountability may do good or harm according to this context. In a learner focused system there needs to be attention to the level of granularity in the way pupils are aggregated into groups and matched with professionally appropriate pedagogy and other supports. Poor performing Maori pupils underachieve for many of the same reasons that poor performing non-Maori pupils do - socio-economic status, family background, health status, peer pressure etc. But the explanation for poor performance is secondary to the reform of pedagogical and other supports aimed at improving performance. Some of the changes will be the same as those required for poor performing students generally, while others will necessarily and appropriately be specific to the circumstances of Maori pupils who are underachieving.

The Treasury-MoEd material provided lists some factors associated with Maori under-achievement (engagement with other social services, leaving school younger, parking lots, quiet journeys). A lot of other and advice e.g Mason Durie's writings, point to the need for responses that are grounded in Maori culture and circumstances. It is evident that implementing this more successfully must recognise that engagement with Maori communities is always local as there are important differences not only between iwi but within them. Also the circumstance of urban Maori must be taken account of. This cautions against the use of accountabilities beyond those of a general nature that require a nationwide response. The answers are going to emerge from engagement with local families, hapu and iwi about how best to design programs for higher achievement by local Maori pupils.

Small details really matter. For example consultations by the Productivity Commission with an urban Maori authority cited an incident in which the authority became concerned about how young Maori pupils were falling behind in mathematics. They put this partly down to the fact that the parents of these children were not helping them with their maths. They approached the education authorities for support for after-school coaching for these children and were refused. The story illustrates that engagement with Maori communities is not all about Maori culture as such and includes practical responses to specific educational needs. The story also illustrates that the current educational system is encrusted with regulations and constraints which can prohibit it from doing some common sense things. The answer, as implied in the terms of reference, is reform to some rigid and counter-

¹⁰ Alistair McCulloch (2009) The student as co-producer: learning from public administration about the student-university relationship, *Studies in Higher Education*, 34:2, 171-183, <http://dx.doi.org/10.1080/03075070802562857>

productive regulations, controls and guidelines to better encourage and enable teaching professionals and school boards to respond more effectively to the individual needs of the pupils and Maori pupils in particular.

Revisions to accountability settings can support greater flexibility and more reliance on seasoned professionals to use their initiative but some of the rigidities in the system are rooted in the rigid and overdesigned HR system. The ferocious opposition of the education unions to alternative governance arrangements is similarly unhelpful to the cause of greater focus on learners.

There are a variety of promising initiatives to address Maori educational underachievement. The Tuhoe, for example, a working with the Ministry to adapt the concept of a Community of Learning to facilitate the tribe's objective to put their young people onto pathways of learning from early childhood through to young adulthood that will help the individuals to succeed in life and also support the long-term developmental aspirations of the Tuhoe. Part of the plan is to revise the curriculum to incorporate learning within the context of the history and knowledge of the people and to apply general educational achievement to the local circumstances and environment. In other words they seek to adapt the state schools in the area to better support their development.

Another approach elsewhere is the development of Kura Houa partnership schools that are tailored specifically to the preferences and needs of some Maori. These have greater freedom than state schools over teaching methods and HR. The accountability for these is through the fixed term contract. Over time as experience with these schools evolves it will signal areas where the rigidities in the state school system are holding back innovation to respond more effectively to the needs of Maori pupils.

The Tuhoe initiative is very much a reflection of their Treaty settlement. This signals a different era emerging in which post settlement iwi seek more control through governance, finance and other means of the education of the children. The revised education act should provide a framework for promoting more productive engagement with Maori about education.

Conclusions

This report concludes by returning to the questions in the terms of reference and responding to them on the basis of the information and arguments presented above.

- 1) Accountability settings are just one element of the relationships in a network covering the school system and which exhibits some characteristics of a complex adaptive system. What drives emergent system outcomes cannot be predicted by reductionist analysis of each of the nodes in the network and assuming that the whole is the sum of its parts. Outcomes emerge from dynamic nonlinear feedbacks rooted in the interplay of the nodes in the network. Although there are strong vertical lines of force through the system emanating from state funding and control, they do not appear to be sufficient to enable the state to drive the system to its preferred outcomes. There are too many autonomous agents whose behaviours are crucial to outcomes and are only partly responsive to the levers in the hands of the state. To deny this is by implication to infer that the state sought the system performance we see, including the long tails.
- 2) Even within a single organisation, accountabilities are only one feature of the organisational architecture consisting of vision, strategy, values, governance, business models, roles, delegations, resources, information, monitoring, reporting, evaluation and innovation. This makes it very difficult to pass judgement on particular accountabilities outside this context in which they are embedded. For a complex system with many organisations the difficulty of

such judgements are magnified. Apart from a few near universal accountabilities for financial control, HR delegations etc, whether a particular accountability is a net benefit to system performance requires its assessment in context.

- 3) Within this context, experience in health and elsewhere suggests that where service improvement requires the collaboration of independent parties the route to better outcomes starts with building those “alliances” in getting agreement between frontline service providers about the “pathways” along which they will pass their clients in order to remove fragmentation and maximise client satisfaction and outcome. The accountabilities in this context follow on from these agreements and generally cannot be imposed from the outside from the start. These accountabilities once they are agreed are nonetheless essential.
- 4) In a flipped school system it follows that the accountabilities for success in a learner focused system would evolve, from the present somewhat hierarchal arrangement, into mutual accountabilities arising from agreements by frontline professionals and those supporting them about how they are going to improve the learning outcomes of the students. By way of comparison, the more successful integrated service delivery models in health engage clinical professionals in agreements about best clinical practices. In the IPIF model there is a dual accountability system in which clinicians accumulate and diffuse knowledge of good medical practices and patient pathways, whereas the state’s accountability system controls objectives, high level priorities, resources and some governance and accountability functions. It is notable that the health system has more senior clinical people with clear professional advisory responsibilities in high places than appears to be the case in education. Creating such positions in education and charging these people with professional leadership would seem essential for a flipped education system. Capability problems across the large number of schools need to be addressed with much greater support than seems to be the case and through accountability for that support being directly relevant to the frontline, rather than just the State’s necessary chain of accountability to Parliament.
- 5) In light of the above, the purposes of the accountability system should be to motivate and support sustained system wide improvement in the services delivered in the classrooms and surroundings. And also to manage poor performance by teachers on the frontline. The accountabilities have to be an element of a system wide architecture aimed at performance improvement. Many of these accountabilities will be multi-party, horizontal and downward from the perspective of the traditional hierarchy. Many will be mutual accountabilities within groups allied to achieve agreed group goals.
- 6) The ‘key accountabilities’ therefore are those that have the largest and most positive effects, not only on teachers, principals and schools but also on pupils by reference to a coproduction concept. Such accountabilities better tailor the teaching and other support for students to their needs and also their goals and commitments. At present the accountabilities for targeted average measures of achievement have distorting effects on the prioritisation of teacher and school efforts. Removal of these distortions will be assisted by recourse to rapidly improving data and analytics that can guide improvements in prioritisation. Key accountabilities also need to be largely by agreement to avoid the agents being swamped and unsupported in meeting poorly researched targets. In this vein, officials should not only be accountable upwards to their ministers but sideways to their peers and downwards to those over which they have oversight. This is essential to the performance of

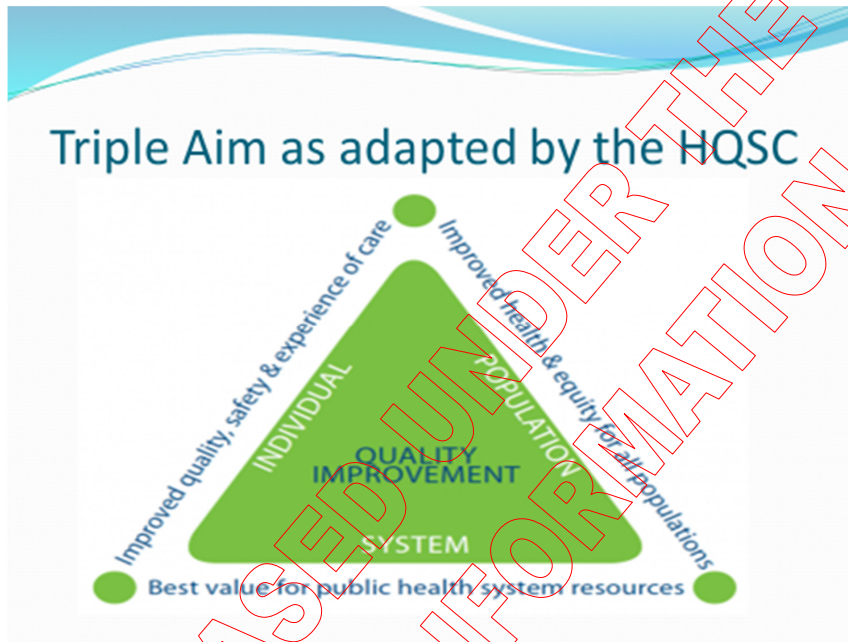
a complex systems whose emergent outcomes are largely driven by these networks of relationships.

- 7) How can accountabilities specifically help to improve learning outcomes of children at risk of not achieving? The Treasury papers note that poor educational attainment is typically associated with other social service needs. In common with other social services, educational outcomes are at risk for children who typically have challenges of a personal nature or in their families or communities. Addressing these challenges requires responsiveness to the specific circumstances of these children either individually or in groups, depending on the circumstances. For example, the Manaiaikalani Trust has evidence that unstable housing has a large negative effect on their programs to lift educational attainment of disadvantaged children to their peers. Schools, principals and teachers cannot as a rule play a part in dealing with such housing issues, whereas this particular trust has the capability to do so. This is a typical example of the broad conclusion in the Productivity Commission report on social services that addressing the long tails requires integrated service delivery through organisations that are suited to the purpose. The network relationships and associated accountabilities must therefore incorporate teachers and schools within institutional structures that can both apply their expertise when the situation of an underachieving student calls for that expertise, and also call on other organisations and sources of support when the situation calls for action that is not within the purview of education as we currently conceive it. Work currently under way to develop the investment approach to social services provides some insights on how this kind of response might be set up.
- 8) Regarding the question in the TOR about hard versus soft levers of accountability, what is the difference? Presumably a hard accountability is a requirement to meet a measurable target whereas a soft one might be something hard to measure like building capability for future service delivery or to establishing a culture that promotes innovation. In the light of the arguments in this report, sustained performance improvement depends on attention to the critical success factors that are necessary for this – regardless of how measurable they are. But again, who is accountable to whom? The hard accountabilities tend to be performance targets and budget constraints set by ministers, whereas that soft ones tend to apply through a general stewardship obligation that is more likely to be the concern of the Board of Trustees. However it is a necessary requirement of the system as a whole that this accountability is present somewhere in the system and that resources and associated support is reasonably available. ERO is the appropriate body to assist the system in defining and assessing capability school by school.
- 9) Regarding the associated question of the government's role in building capability, it has essentially delegated this to the schools subject to the funding it provides. But perhaps it has over-delegated and could usefully provide more support. The thought of having more leading respected experts in the Ministry or a professional body has been expressed above. Having 2500 schools all trying to lift capability seems unrealistic and will explain some of the long tail. A more focused strategic initiative to lift capability through a package of hard and soft measures is likely to improve the current situation and address the wide variances.
- 10) The TOR ask what the Government should do to create conditions in which professional autonomy and accountability drives continual improvement. There is already considerable professional autonomy in the sense that the curriculum allows schools and teachers freedom to teach it in accordance with what they see as the best pedagogy in the

circumstances. So the question becomes what additional professional autonomy could be allowed for and what effects this would have. As noted above, the school system is very fragmented and there could be more support for continual professional development and stronger professional collegiality to support that. The proposed modification to the Act makes some attempt to do this through reforming the Council into a Competence Authority, although this is largely about hearing complaints about lack of competence. In the health sector there are many professional bodies such as the colleges, which take a wider view of the standards and good practices expected of members. In education the unions seem to be the strongest voices from the teaching profession and it is desirable that other voices that focus on professional development are strengthened. Perhaps a reformed Council can do this. The quest for a learner focused system will be undermined unless there is reform of the HR systems to strengthen the incentives and rewards for teachers to lift their capabilities and be more discriminating between levels of competence.

- 11) Finally, addressing the over-representation of Maori in the long tails of education, the place of accountabilities within a system that improves outcomes for Maori pupils should be derived from the institutional architecture that has been chosen for the purpose. A particular accountability may do good or harm according to this context. The NCEA target under the BPS system has arguably been harmful. Like any achievement based on a proportion of a population it motivates providers to focus their attention on the people who are just below the line and can get over it without too much effort. This turns their attention away from the long tails. A learner centred education system which is based on incremental achievement, student by student is much to be preferred - especially for Maori and others in the long tails. New data analytics are available to make this kind of system more feasible than it has been. From this could flow more refined approach to objectives and accountabilities for performance.

Annex 1: Integrated Performance and Improvement Framework for the health sector



	Triple Aim Dimensions			Incentives
Breakthrough	Improved quality, safety and experience of care	Improved health and equity for all populations	Best value for public health system resources	Maximum freedom to innovate, flexible accountability framework
	<ul style="list-style-type: none"> Measures not predetermined Measures apply equally to Practices, PHOs and DHBs, reflecting the goals of programme innovation 			
Excellence (Higher thresholds)	System level measures in common across practices, PHOs, DHBs	Contributory measures, for each of practices, PHOs, DHBs	Contributory measures, for each of practices, PHOs, DHBs	Financial Reputational Increased influence Increased ability to lead innovation
Improvement	System level measures in common across practices, PHOs, DHBs	Contributory measures, for each of practices, PHOs, DHBs	Contributory measures, for each of practices, PHOs, DHBs	Financial Reputational
Entry	DHB Accountability Framework Quality Accounts PHO Minimum Requirements RNZCGP Foundation Standards			Sanction