Towards an optimal model for schooling improvement

Building Evaluative Capability in Schooling Improvement Project

Helen Timperley, Stuart McNaughton, Mei Lai, Margie Hohepa, Judy Parr and Rachel Dingle

April 2009

University of Auckland

New Zealand Council for Educational Research

Overview

This position paper provides a broad outline of an optimal model for schooling improvement for the New Zealand context. Effectiveness is defined in this model as the sustained achievement of student outcomes valued by the community. The model is based on the premise that teachers have the greatest system influence on student outcomes but that they cannot achieve such improvement alone. All layers of the system—teacher, schools, clusters and the Ministry of Education (MOE)—have a shared responsibility and accountability for better outcomes.

Core Concepts

We propose three core concepts as basic to schooling improvement. The first of these is *capability*, which means having the knowledge and skills to carry out the specific roles, responsibilities and actions required to improve student outcomes. An important dimension of capability is evaluative capability, which allows participants in schooling improvement to formulate targets, develop effective interventions and have evidence-informed systems to judge progress towards the targets.

> The core concepts of capability, managed interdependence and having relationships of trust and challenge are described on pages 6-9.

The second core concept is managed interdependence. For schooling improvement to be effective, each level of the education system needs to have complementary and mutually informed relationships with outside agencies or with other levels of the system.

The third core concept involves having relationships that are trusting yet challenging. Trust involves personal respect, integrity and carrying out mutual agreements. An understanding of mutual accountabilities is essential to high trust. Trusting relationships, however, do not necessarily result in improvement unless they are accompanied by task-focused challenges that disrupt current routines and practices, and result in doing something different.

Cycles of inquiry

Central to developing the knowledge and skills for high levels of evaluative capability is engagement in cycles of evidence-informed inquiry with a focus on students and their learning. The cycles not only build evaluative capability, but also the knowledge and skills needed to improve practice through the inquiry. The case is made that such inquiry needs to occur at all levels of the system, including the Ministry of Education.

> Cycles of inquiry to build teacher and leader knowledge are outlined on pages 9 –12.

Progressions

Knowing how to undertake the difficult business of schooling improvement usually takes time as schools develop specific capabilities as they work through the process. Schools don't suddenly change from a "cruising" position to one systematically focused on raising student achievement. Thus, three levels of progression (basic, middle/mixed and integrated) are identified on pages 14–17 that are designed to describe a typical pattern of progress. These progressions allow those involved in schooling improvement to think about where they fit on each dimension and what evidence supports their decision. Once located, they are able to see how to make further progress.

Domains of knowledge and action that contribute to schooling improvement

Many domains of knowledge and action contribute to the development of capability and interdependence through the progressions. The four described include the evidence base on which to make judgements of progress; theories for improvement and sustainability that underpin schooling improvement activities; professional learning communities as the forum for much professional learning; and planning for building evaluative capability.

Important domains of knowledge for building capability are outlined on pages 18–25.

The evidence base needs to include achievement outcomes valued for students, together with evidence of those practices and processes thought to have an impact on such outcomes; in particular, those related to teaching and leadership.

A theory for improvement is just a set of linked ideas about how to improve something. In essence, it includes a definition of the problem the theory is designed to address, proposed solutions to address it, a rationale for the solutions, targets against which progress can be judged, and ways to monitor such progress.

Professional learning communities can form a powerful force for change. Learning in such a community involves two facets—the content for discussion and the interpersonal interactions among members of the community. For optimal learning, the community must be strong in both.

Planning to build evaluative capability in schooling improvement requires more than the formulation of a schooling improvement plan. It also requires that those involved understand the principles underpinning effective planning processes so that, in the future, they are able to develop their own plans in association with, or independently of, external expertise.

Māori medium education (MME) is a context of particular significance in New Zealand. We have taken the position that an optimal model for

Issues concerning Māori medium education are outlined on page 26.

schooling improvement must include opportunities to enhance outcomes for those students receiving their education in Māori medium settings. Inclusive frameworks require the development of a theory for improvement that ensures that any improvement focus does not work in opposition to Māori language and cultural development, which is fundamental to these contexts.

The paper concludes with our theory of learning from the position papers. Pieces of paper are poor mechanisms for change. The ideas expressed here must become part of the embedded routines of schooling clusters, with accompanying opportunities to learn the underlying knowledge and skills crucial to the ongoing work of schooling improvement in New Zealand.

Introduction

In this position paper we present an initial model that represents our thinking about effective forms of schooling improvement. The model can apply equally to any schooling development work where the aim is to accelerate student achievement, particularly for those students who are not reaching the standards expected for their age or years of instruction in a particular language. It applies to both individual and clusters of schools seeking to improve student outcomes. It equally applies to the Ministry of Education (MOE) improving the way they support schools to improve.

> Schooling improvement is the core business of all schools.

There are two purposes behind the model. One is to present our thinking about building sustainable evaluative capability in schools that has guided our research and evaluation. The second purpose is to communicate our thinking to the MOE and the schools engaged in schooling improvement work. We are anticipating that the model will provide a basis for critique of our thinking and, subsequently, a basis for critique and development of schooling improvement practices.

The model is based on the premise that teachers have the greatest system leverage to make a difference to student achievement (Alton-Lee, 2003). We also know that teachers cannot achieve this on their own, but that the MOE, school leaders and the quality of the professional learning offered to teachers support a chain of influence on teaching and student learning. All layers of the system have a shared responsibility and accountability for improved outcomes for students, particularly those achieving less well than others.

This initial paper presents an overview of how this might happen, with many of the ideas developed further in additional papers related to specific components. Schooling improvement is a nested system in which specific activities occur at various levels. Currently these levels include the classroom, the school, clusters of schools and the MOE. Student learning and development is influenced over time by each level and the expertise invoked to promote learning at that level. In any schooling improvement model, improving classroom instruction and leadership practices in schools is fundamental to success. The presence of a cluster may or may not be integral to that success. It is not the structure that matters. Rather it is the quality of what happens within each school as a result of participating in a particular structure that matters.

Within any schooling improvement approach, there are features that are of particular significance to that context. In New Zealand, a key feature is the presence of Māori medium education (MME). How Māori medium is located within schooling improvement has implications across all levels of the system. This location is also a critical predictor for the success of schooling improvement efforts in Māori medium contexts. The proposed papers also include one focusing on MME.

The model we present is an idealised one and assumes that there are developmental progressions towards what we would see as an optimal form of schooling improvement. There are various sequences that clusters or other participants might take. Their development is not stage-like and unitary; in other words, there is not a fixed single pathway. Features of the likely developmental progressions are described further in a subsequent section of this paper.

The model

At the core of our model are three core concepts: increasing capability (including evaluative capability); encouraging schools to self-manage their interdependence with other agencies; and building relationships throughout the system that are trusting yet challenging. As schools and other participating groups develop these attributes, they move through a series of developmental progressions. We identify these and analyse typical activity at each progression. Four domains of knowledge and action make crucial contributions to the process. These are: the evidence base; theories for improvement; professional learning communities; and planning for building evaluative capability.

The remainder of this paper will expand on and explain each element of the model, why we believe they are important and how they work together.

Core Concepts

In this section we elaborate on what we mean by the core concepts of capability, managed interdependence, and trust and challenging relationships.

Capability

Capability means having knowledge and skills to carry out the specific roles, responsibilities and actions needed to improve student outcomes. Capability can range from having little or no capability in each domain through to having high capability.

> Capability = knowledge and skills to carry out specific roles and responsibilities.

One aspect of capability is that the teachers and leaders in the schools within a cluster need a certain level of knowledge about instructional practices. This knowledge is not purely programmatic or technical. Teachers are professional experts who are deeply knowledgeable about what they do, how they do it and why they do it. Their knowledge and skills are about particular practices such as literacy content and pedagogy, and using particular forms of guidance and assessment tools in particular settings. They have articulated knowledge and understanding of the domain of teaching, and of children and their learning. They understand how to teach effectively; they have strategic practices that are versatile and adaptable; and, finally, they show a keen awareness of the effectiveness of their practices both through reflection on and regulation of

these practices (McNaughton & Lai, 2009). However, our view of the research literature suggests that having technical capability to implement a pre-specified programme does not guarantee sustainability of schooling improvement initiatives and indeed may limit further problem-solving.

- Another apsect of capability in schooling improvement relates to developing knowledge and skills to evaluate the impact of change and to make appropriate adjustments when the impact is less than aimed for. This aspect is called *evaluative capability* and it is central to making ongoing improvements to professional practice and student learning. Evaluating this impact, diagnosing possible ways to enhance it and making appropriate adjustments requires detailed understanding of students' learning needs.

Identifying and engaging in the type of professional learning required to improve outcomes for students, together with an ongoing evaluation of its effectiveness in terms of changes in classroom practice and in student learning, is part of evaluative capability.

Managed interdependence

Interdependence refers to the extent to which having the capability described above has allowed the school to come to be relatively independent in the sense of management and control. It also involves complementary and mutually informed relationships with outside agencies, or with other levels of the system. These relationships typically range from dependence on external agencies to having self-managed interdependence. It is possible for a cluster to have the necessary capability to improve student outcomes because they are collaborating with an external agent, but they may be dependent on that external agent. For example, an external agent may design and carry out each of the steps involved in collecting achievement evidence. It is also possible for one level of the system to become dependent on another level for its roles, responsibilities, knowledge and action. For example, part of a school's role is to resource the professional learning of its teachers, but the school could become dependent on the MOE for that resource.

The research literature (Bandura,1995) suggests that high dependence on external agents and/or other levels in the system may undermine effectiveness in terms of student learning and in terms of developing evaluative capability. That is, the development of capability may be undermined if the external agent or other level of the system holds the critical information on achievement and instructional practice and individual teachers are dependent on access to the agent for that information. Similarly, over time, too great a dependence can lead to a kind of "learned helplessness", where teachers and schools do not develop the capacity to make their own inquiries into whether their teaching practice is working adequately.

While in general schools need to achieve a level of independence in order to develop effectiveness and evaluative capability, it is important that some form of interdependence be maintained. It is unrealistic to assume that schools could develop full autonomy in each

domain. But, more importantly, having interdependent relationships, particularly with those with specialist expertise, is necessary to developing expertise within any one school community in ongoing ways. This becomes the responsibility of the school to manage because as old problems are solved and new challenges emerge, new relationships involving new expertise need to be forged.

Trust and challenge

The interpersonal dynamics in any schooling improvement project influence success because the process of improving capability and developing interdependence is usually associated with a sense of vulnerability. Tony Bryk and colleagues (Bryk, Sebring, Allenswroth, Luppescu & Easton, 2007) in the United States identify the importance of having a base level of trust, forged through day-to-day social exchanges, as fundamental to success. Trust helps to develop buy-in, together with the motivation and deep engagement to undertake the difficult work of schooling improvement. These authors define trust in terms of personal respect, integrity and competence in the execution of basic responsibilities related to particular roles. This definition goes beyond the idea of trust as uncritical support of one another. It has a strong element of mutual accountability. Trust is based on perceptions of personal integrity that those involved can be trusted to keep their word and carry out mutual agreements. The specifics of how this looks in any particular situation will differ according to the cultural and contextual conditions.

> Trust involves personal respect, integrity and carrying out mutual agreements.
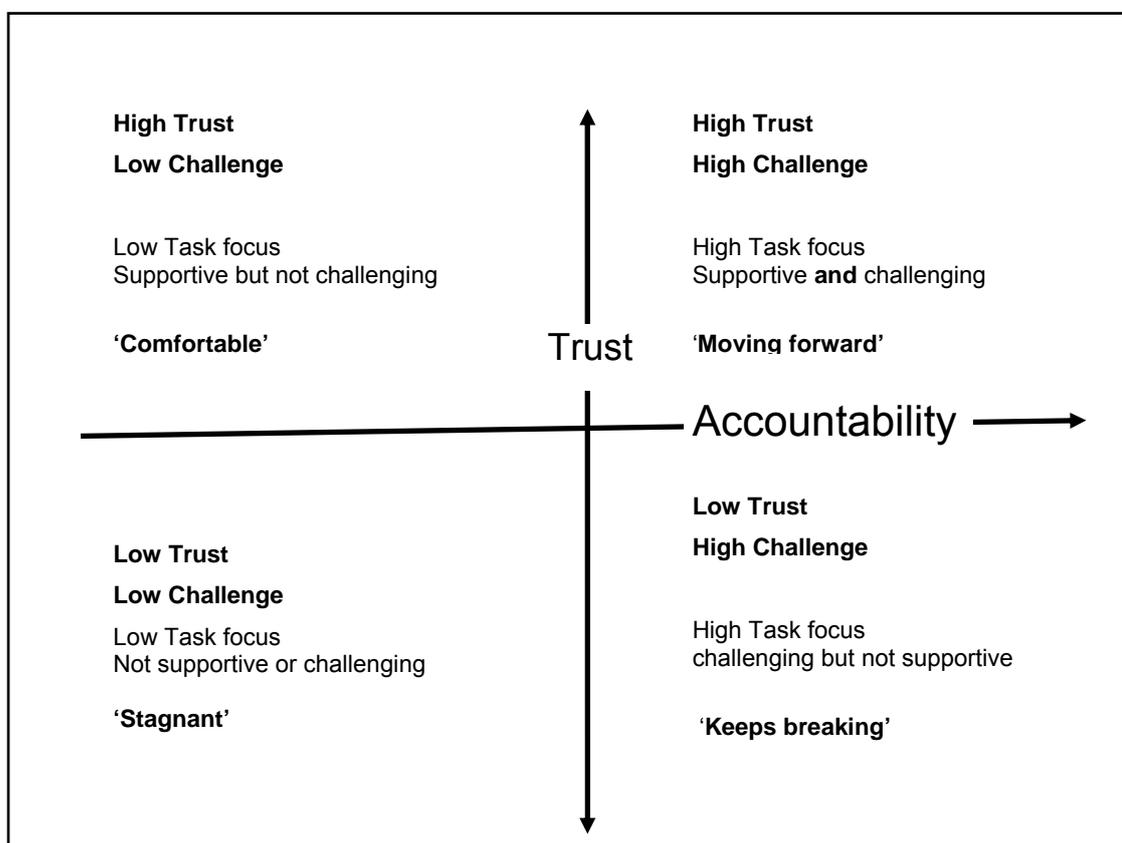
These authors also identified that the development of trust does not have a direct impact on student achievement. This impact comes from the mutual learning that happens as a result. A dynamic of uncritical support, where all statements are taken as equally valid, does not promote such learning. Rather, the social exchanges need to develop into challenge of each other's claims, the beliefs underpinning them and the evidence on which they are based. Trust comes from a perception that such challenges will be undertaken respectfully and in a supportive way. For example, when identifying possible causes of student underachievement, it is important that the ideas put forward are examined for their worth and rigorously debated. Otherwise, it is quite possible that the proposed actions to address the causes will not do so, and professional and student time will be wasted. High trust and high challenge requires a strong task focus on solving the identified problems in a supportive environment. Paul Wright from the Manurewa Enhancement Initiative adapted a diagram from Hopkins, Ainscow and West (1994) *School Improvement in an Era of Change*, which summarised how he saw the relationship between trust and challenge (Figure 1). He identified that a strong task focus with high trust and high challenge is needed for the schools involved in the cluster to move forward. Any other combination fails to achieve the aim of

> Challenge requires rigorous debate of the ideas put forward and the evidence that underpins them.

improving student outcomes. In the inventory phase of the Building Evaluative Capability in Schooling Improvement Project, such challenge was rarely evident in cluster meetings (Timperley, Parr, Hohepa, Le Fevre, Lai, Dingle, & Schagen, 2008).

Figure 1 Trust and Challenge

High Trust **Low Challenge** Low Task focus Supportive but not challenging **'Comfortable'**	**High Trust** **High Challenge** High Task focus Supportive **and** challenging '**Moving forward'**
Low Trust **Low Challenge** Low Task focus Not supportive or challenging **'Stagnant'**	**Low Trust** **High Challenge** High Task focus challenging but not supportive '**Keeps breaking'**

Trust — Accountability →

Source: Paul Wright, 2008

Building professional knowledge and evaluative capability through cycles of inquiry

In our view, engagement in evidence-informed inquiry is the best way to develop evaluative capability and productive interdependence. Such inquiry involves working through ongoing cycles that explicitly build knowledge of assessment, curriculum and instruction (Timperley, 2008). Cycles of inquiry that improve student outcomes by changing teaching practice began by engaging teachers in examining evidence of student learning. Gaps between outcomes that have been identified as valued by the community and current achievement then form the rationale for teachers to engage in professional learning (see Figure 2).

Figure 2 Building knowledge through cycles of teacher professional inquiry
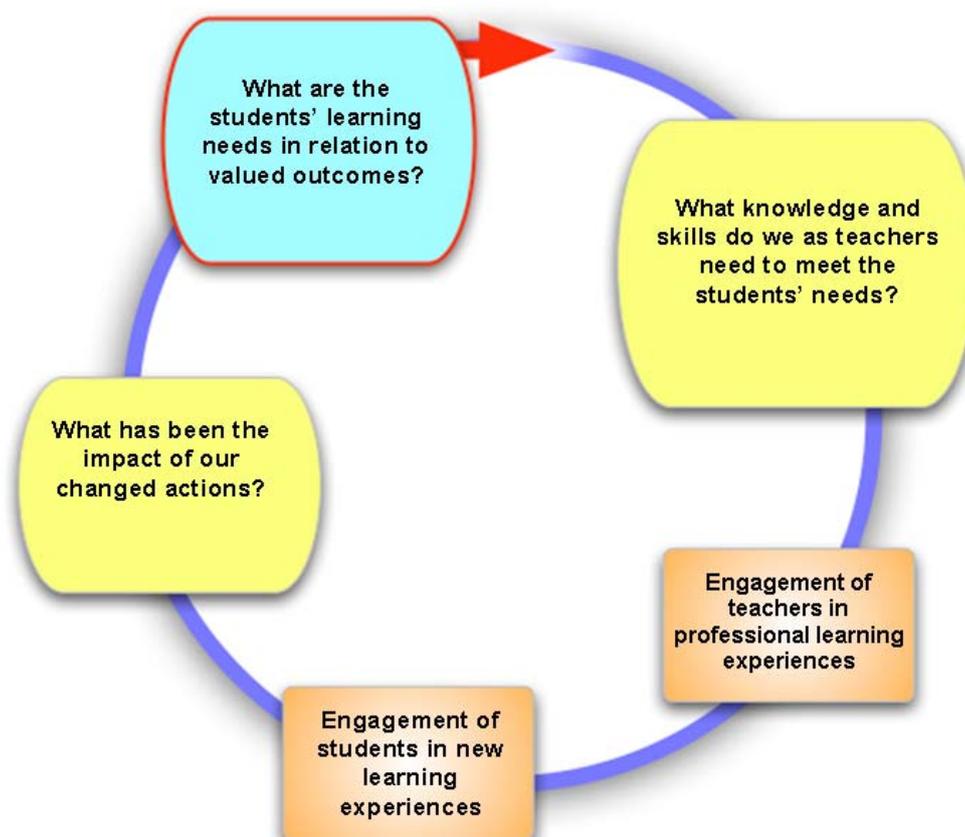
The participating teachers, with the assistance of relevant experts, are then asked to consider what they need to learn. They begin this process by asking questions about how they have contributed to existing student outcomes. Additional questions include what they already know that they can use and what they need to learn and do to improve student learning. From the answers to these questions, teachers co-construct with relevant experts what it is they specifically need to learn about teaching practices and assessment that is relevant to the student learning needs they are trying to address. Their engagement in professional learning should always be underpinned by knowledge of their students' learning needs and a focus on how to improve instruction to meet those learning needs. As they engage in professional learning, teachers then construct new learning experiences for their students. The cycle is completed by seeking evidence about changes made to teaching practices and the impact of those practices on student learning. The links between any particular teaching activity and student learning are complex, so this last step is critical in determining whether new practice is having the anticipated impact. From this evidence new cycles of inquiry are begun.

Although most successful schooling improvement initiatives involve professional learning for teachers because it is they who have the greatest impact, we cannot expect teachers alone to solve challenging achievement problems. The organisational conditions in which teachers work and the learning opportunities provided by their school leaders are critically important. Effective leadership is identified consistently in both the schooling improvement and school

effectiveness literature as central to improving student outcomes (*International Handbook of School Effectiveness and School Improvement*). A meta-analysis undertaken by Robinson, Hohepa and Lloyd (in press) identified that the most powerful thing school leaders can do to improve student outcomes is to promote and participate in the professional development of their teachers. No doubt this finding would apply to cluster leaders in situations where schools are clustered for schooling improvement activities. Cluster and school leaders, therefore, also need to participate in professional learning cycles about how to lead schooling improvement initiatives in the same way as the teachers. Figure 3 shows how school and cluster leaders need to engage in their own cycles of inquiry to build their knowledge to become more effective schooling improvement leaders. The impact of one layer of the system on another means that each is interdependent with the other.

Figure 3

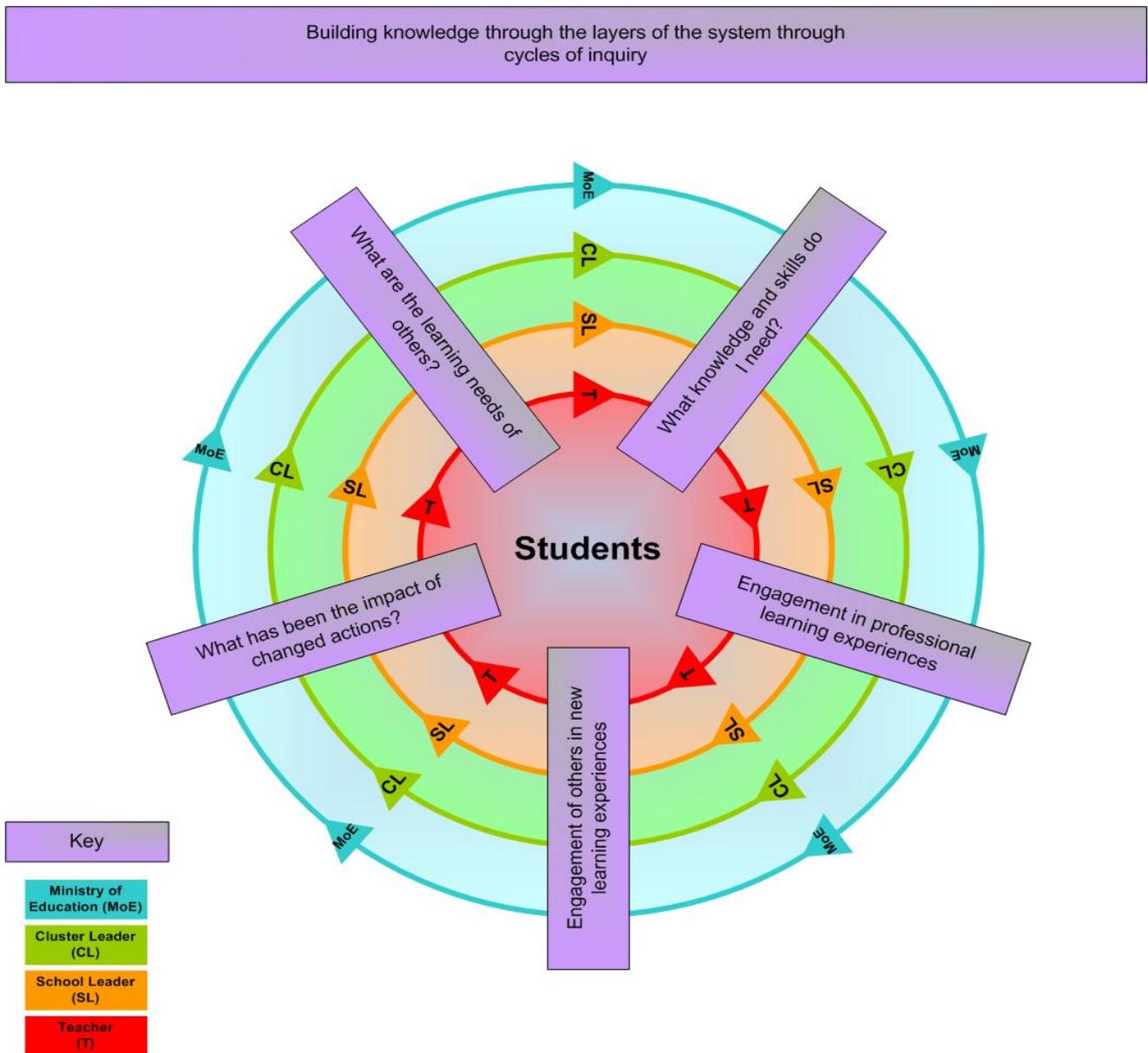Building knowledge through the layers of the system through cycles of inquiry

Figure 3 includes a learning and knowledge-building cycle for the MOE on the outermost ring. Leading schooling improvement is a complex task and the leadership of the MOE is critical to success because MOE co-ordinators have the opportunity to develop in-depth understandings of effective interventions and the circumstances under which they are likely to be successful across schooling improvement interventions. We suggest that MOE co-ordinators need to assist clusters to identify effective interventions from relevant research and contextualise these interventions according to the particular circumstances in the clusters . In the initial stages, there is likely to be a high level of dependence in making such choices, with greater independence developing over time. Other MOE roles might include challenging cluster decisions that may not be the most effective in terms of promoting accelerated student learning. If MOE co-ordinators are to maximise their effectiveness in this role, then they too need to engage in inquiry cycles that focus on meeting the needs of the schools and groups within the cluster.

Developmental progressions

Engagement in the inquiry cycle is fundamental to the development of evaluative capability. With new clusters, the type of engagement is often superficial and occurs at a high level of generality. For example, the process of identifying student learning needs might involve comparing cluster averages on a single assessment with national averages on that assessment at one point in time. As clusters develop their evaluative capability, analyses become more nuanced and might assess targeted subgroups and progress over time. This section outlines how clusters may progress in

> It is important to think about where your school(s) fit on the progressions and what evidence supports your decisions.

their evaluative capability as they engage in cycles of inquiry. The cycle illustrated in Figure 2 is used as the organising framework. The specifics of these cycles look different at different levels of the education system, but our focus in this paper is on the cluster. The cycles and associated progressions may equally be applied to the school and classroom levels and this will be illustrated in a later paper. The operations of the MOE and the clusters, of course, both influence and are influenced by the operations of the schools. No part of the system is independent of the others. Indeed, research by Katz, Earl and Ben Jafaar (in press) indicates that clusters have only an indirect impact on schools. Participating in a cluster only adds value when the schools themselves are functioning effectively.

In these progressions we outline three possible points on a continuum of evaluative capability. We know from research into schooling improvement that optimal models are effective only if those using them can see where their current practice sits and have some knowledge of what they need to do to improve. Knowing has a developmental history and schools cannot jump from a basic to integrated level without developing the systems, understandings and processes along the way. For this reason we have briefly portrayed typical characteristics of

a cluster working according to three progressions: basic, middle/mixed and integrated. In doing so, we are not suggesting that development is stage-like or unitary, or that a cluster would neatly fall into one of the progressions. Rather, the descriptions should be seen by the schools and clusters as an approximate fit. Many schools and clusters at their beginning stages may not demonstrate practices consistent with even the basic point on the continuum.

Bullet point descriptors are used as illustrative examples of how each dimension of the inquiry cycle (Figure 2) might look at the basic, middle/mixed and integrated progressions for cluster leadership on the core concepts of capability, interdependence and trust. The descriptors apply at the cluster leadership level, with similar ones for the MOE given in Appendix 1. The same principles also apply at the individual school leadership, teaching and student level, and the descriptors can be readily translated to these levels.

Table 1: Progressions of evaluative capability for cluster leaders at each step of the inquiry learning cycle

Dimension 1: Identifying valued outcomes and student learning needs of schools and groups within the cluster

Progression	Typical cluster activity	Capability and Interdependence
Basic	- National levels of student achievement assumed to be valued and appropriate for all contexts - Students assessed using nationally benchmarked assessment - Results entered into database and basic analysis undertaken—calculation of averages, results of improvement graphed over baseline over a year - Cluster and school (anonymous) averages compared with national averages and area of identified need (e.g., reading comprehension) - Intervention decisions made about focus at general level (e.g., "our students are low on writing") - School leaders discuss results and report to staff	- Low capability because internally limited and uncritically reliant on internal resources - High dependence, reliant on MOE or external expertise to drive cluster
Middle/mixed	- National levels of student achievement are assumed to be those valued, with some cognisance given to the language of instruction (e.g., MME) - Multiple sources of evidence collected but incomplete for all schools in the cluster - Different sources of evidence used as a basis to evaluate and plan cluster activities	- Mixed capability—some emerging skills and resources to collect and use appropriate sources of evidence - Not aware of need to be *mutually* informative through relationships (e.g., with MOE providers)
Integrated	- Shared understanding of what is valued is achieved among whole school communities (e.g., where MME is part of a cluster, there is shared understanding of MME-valued outcomes) - Assessments:<ol type="a"> - aligned with the valued outcomes - standardised procedures for administration and data entry used - appropriate for the language of instruction - Questions asked of the data arise from identified focus - Analyses answer the questions asked, with triangulation across multiple data sources and further data requirements identified - Explanations sought for patterns of progress include whole school community - Specific decisions about intervention focus made (e.g., students have limited ability to use the resource of language for the purpose intended) - All involved discuss and understand their contribution	High capability and well-managed interdependence

Dimension 2: Identifying professional learning needs—leaders and teachers

Progression	Typical cluster activity	Capability and interdependence
Basic	• Professional development (PD) in area of focus determined for teachers and assumed to be generic (e.g., PD in writing) • Leaders know the work plan and organise their schools to allow it to happen (e.g., giving release time) • Neither leaders' nor teachers' learning needs determined by evidence and evidence not discussed	• Imported foci which may or may not relate to actual needs
Middle/mixed	• Student achievement information used to assess teachers' learning needs but may not be acted on • Leaders' training undertaken in specific leadership practices not integrated into leading the schooling improvement plan	• Low capability in using evidence to directly inform practice through leadership • Little mutual influence at the level of leadership (e.g., between leaders and external PD agencies)
Integrated	• Student data used to identify what has and has not been taught well and what next gaps appear to be • Leaders' and teachers' (including MME) professional learning needs determined through systematic observation or knowledge assessment by those with established expertise • Evidence discussed with leaders and teachers and respective views tested • Interventions based on a deep understanding of how professionals learn, with particular attention paid to the underlying coherence of different cluster initiatives	• High capability and well-managed interdependence

Dimension 3: Identifying own learning needs—cluster leaders

Progression	Typical cluster activity	Capability and interdependence
Basic	• Not explicitly identified for cluster leaders	• Leadership of cluster assumed to be effective
Middle/mixed	• Self-identified and not based on specific or systematic evidence for cluster leaders	• Still dependent on needs defined primarily by general expectations or external agencies (e.g., MOE)
Integrated	• Complex challenge of improving schools explicitly recognised, with evidence about effectiveness sought in a range of areas and learning goals established • Sufficient knowledge base developed to make strategic resourcing decisions and supportive policies and actions	• High capability and well-managed interdependence

Dimension 4: Engagement in own professional learning to deepen knowledge and refine skills

Progression	Typical cluster activity	Capability and interdependence
Basic	• External or internal provider runs workshops and coaching sessions with teachers in and out of class, so no further direct leadership development considered relevant.	▪ Reliant on other expertise for the design of the PD
Middle/mixed	▪ No systematic engagement perceived to be necessary	▪ Little capability to trigger and develop internal processes of learning
Integrated	▪ Leadership development focused on purpose of successful implementation of schooling improvement focus across cluster schools including MME ▪ Leaders engage with or lead other professional learning experiences	▪ High capability and well-managed interdependence

Dimension 5: Assessment of impact

Progression	Typical cluster activity	Capability and interdependence
Basic	▪ Students reassessed and averages compared with previous averages and national averages at school and cluster levels ▪ No other data systematically monitored to identify possible causes of impact or lack of it, although informal observations such as walkthroughs may be used to assess implementation	▪ No systematic evaluation of programme effectiveness
Middle/mixed	▪ Mix of evidence and preference used to establish priorities ▪ Competing claims not challenged	▪ Not yet fully capable of using evidence to identify and solve needs
Integrated	▪ Assessment cycle is formative (daily to annually) using a range of methods appropriate to time span ▪ Possible causes identified, debated and tested ▪ Unintended consequences, such as potential negative impacts, monitored ▪ Interdependence and joint responsibility for monitoring impact taken by students, teachers, school and cluster leaders ▪ Joint evidence-informed inquiry into explanatory factors undertaken ▪ All those involved can answer the questions, Where am I going? How am I going? and Where to next?	▪ High capability and well-managed interdependence

Dimension 6: Re-engaging in the cycle

Progression	Typical cluster activity	Capability and interdependence
Basic	▪ Another cycle re-engaged with minimal adjustment of priorities or operations ▪ Planning based on those aspects of the previous year considered successful and preferred by school leaders	▪ Uncritical dependence on previous practice or use of what's available without reference to proven effectiveness in raising achievement
Middle/mixed	▪ Mix of evidence and preference used to establish new priorities	▪ Some input into future plans but still dependent on external co-ordinator to formulate them
Integrated	▪ Another cycle re-engaged by considering what needs to be adjusted for successful implementation of the plan across schooling provision in the cluster including MME ▪ All schools in the cluster share similar understandings of where they are at and what the most relevant evidence is to consider ▪ Analysis of coherence and contradictions with other cluster activities undertaken ▪ Planning based on evidence and possible explanations, which are publicly shared	▪ High capability and well-managed interdependence

Structural features, assumptions and interpersonal dynamics that influence all dimensions of the cycle

Schools are arranged into clusters with particular interpersonal dynamics that influence all dimensions of the cycle. Those that are often evident at the basic, middle/mixed and integrated levels of the progressions are identified below.

Basic
- Previously developed theory for improvement not considered for revision during the process

- professional learning communities within schools and across schools established for teachers; these are focused on understanding student achievement and implementing external advice

- roles, responsibilities and accountabilities assumed rather than specifically identified, with co-ordination for implementation perceived to be the responsibility of cluster co-ordinators

- a "one size fits all" approach is assumed to be appropriate, with little consideration given to the types of schooling provisions in the cluster (e.g., secondary, MME)

- the development of trust is focused on noncritical support rather than challenge

- existing cluster structures assumed to be appropriate for new work, or new structures added

- cluster relies on MOE or external expertise (dependent) or relies solely on own expertise (independent).

Middle/mixed

- some elements of previously developed theory for improvement are considered for revision but considered in isolation from one another

- professional communities discuss content, but the interpersonal interactions are focused on congeniality rather than learning

- roles and responsibilities identified but no specific reference to mutual accountabilities

- some account taken of the needs of different schooling provisions in the cluster (e.g., secondary and primary), but language of instruction not specifically considered

- development of trust primarily supportive, with elements of challenge (usually in private) emerging

- revision to some cluster structures considered in different cycles but not taken as a whole

.Integrated

- previously developed theory for improvement considered for revision and the elements of the cycle become an integrated theory for improvement

- professional learning communities within schools and across schools strong in both content and interpersonal interactions

- roles, responsibilities and accountabilities identified across school and MOE, with implementation considered a joint responsibility

- account taken of different schooling provisions, including the kaupapa of MME, with shared responsibility to ensure identified needs across schooling types are addressed

- high trust and high challenge evident in cluster interactions

- cluster structures reviewed for both effectiveness and efficiency and to identify any revisions to work programme.

Domains of knowledge and action that contribute to schooling improvement

Many domains of knowledge and action contribute to the development of capability and interdependence through the progressions. In this section, we briefly outline four:

- the evidence base

- theories for improvement

- professional learning communities

- planning to build evaluative capability.

The evidence base

Evaluative capability depends on collecting and using evidence consistent with a cluster's agreed valued outcomes. Given the emphasis on student achievement in schooling improvement, the evidence needs to be

The basics of evaluative capability involve developing an evidence base about student learning (or other valued outcomes), teaching and leadership practices.

about students' learning and classroom instruction in the local context. It needs to be able to be mapped on to, and checked against, national patterns. Local evidence about teaching, learning and leadership practices is needed at the school level for at least two reasons. One is the need to base teaching practice on evidence about teaching and learning drawn from on-the-ground patterns. The second is the need for an individual teacher to be able to use a range of assessments, both formal and informal, within a range of contexts, both formal and informal, to broaden their knowledge of individual students and effectiveness with those students, in order to personalise instruction better. Collection of the evidence needs to be systematic, valid and reliable. Clusters also need to be able to verify that it is adequate.

At the local level, the evidence base for student achievement for schools participating in schooling improvement initiatives is usually reasonably strong. Evidence about teaching and leadership practices is generally less well developed (Timperley et al, 2008).

At a national system level, however, we currently do not have regular, widespread and systematic information about patterns of student progress and achievement in core areas as there is no consistent framework to gauge this. The current work on developing national standards provides an opportunity for the development of a framework that will provide this information. Effectiveness needs to be judged in terms of two criteria. The first is accelerated gains relative to expected growth. This is what enables students to "catch up". The second is the distribution of achievement coming to match national distributions. In other words, the students in schooling improvement schools coming to have the same probability of being in any part of the distribution as any other student nationally.

Theories for improvement

In essence, a theory is just a set of linked ideas to explain something. A theory for improvement is therefore a set of linked ideas about how to improve valued outcomes. Our definition of theory here includes both "formal" theories, for example, research theories about how to improve reading comprehension; and tacit theories, for example, a teacher's ideas on how to improve students' metacognitive skills that are unarticulated but enacted in a lesson. Everyone is a "theorist" in the sense that they have tacit understandings of how to act in situations.

It is important to understand theories for improvement because an individual's or a group's theories will guide practices and will result in cluster members choosing some practices over others. For example, a cluster may choose an intervention that focuses on student engagement rather than classroom instruction because the school decides that improving engagement will result in better student achievement. So, by examining the theories of the schools and the cluster and therefore making any theory explicit, the reasons (and assumptions) for why the cluster has chosen some practices over others are uncovered (Robinson & Lai, 2006). This allows the cluster and its partners to evaluate collectively the
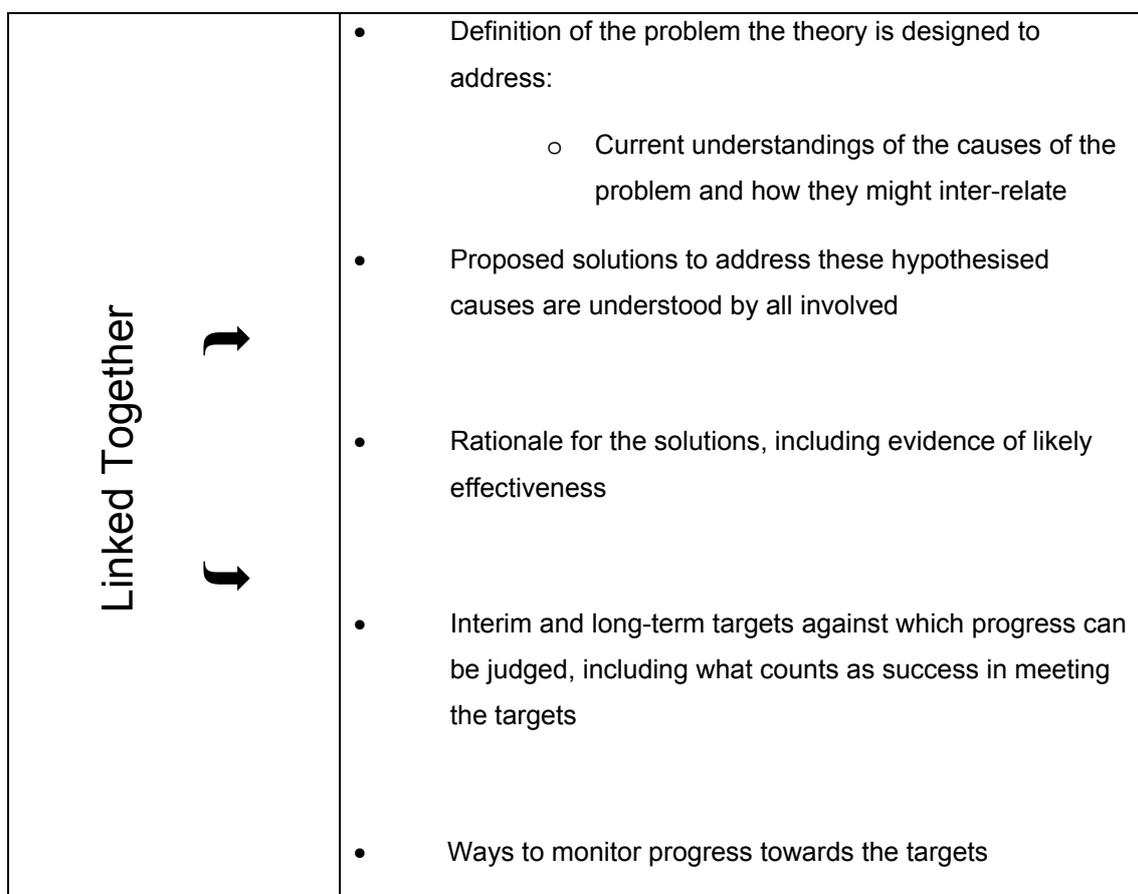
adequacy of these theories (including the reasons and assumptions underpinning them) using agreed evaluation criteria. Evaluating theories is fundamental to effectiveness because not all theories will contribute equally to the cluster's desired outcomes.

To evaluate and adjust theories for improvement, schools, clusters and their partners can check and monitor, for example:

- whether the factual claims underpinning the theories are accurate, for example, "Is it true that children speak only Samoan at home?"

- whether the knowledge base for choosing particular theories is adequate, for example, "Should we really be following the practices recommended by a milestone report that looks only at teacher satisfaction with the PD rather than whether achievement has improved?"

- whether the cluster reasoning and subsequent actions resulted in the cluster's desired outcomes, for example, "Did the PD raise achievement?"

Below is a summary of the authors' current assumptions about the components of a theory for improvement. These need to be grounded in evidence from the cluster, together with relevant evidence from research.

Figure 4 Components of a theory for improvement

Linked Together	
	• Definition of the problem the theory is designed to address: o Current understandings of the causes of the problem and how they might inter-relate • Proposed solutions to address these hypothesised causes are understood by all involved • Rationale for the solutions, including evidence of likely effectiveness • Interim and long-term targets against which progress can be judged, including what counts as success in meeting the targets • Ways to monitor progress towards the targets

A theory for improvement also needs to include sustainability of valued outcomes. Sustainability is a process of organisational learning to improve outcomes already achieved (Lai, McNaughton, Timperley & Hsiao, 2009). As such, our theory of sustainability is that it is an ongoing process that follows on from initial improvements. For this reason, sustainability is not something that is planned for after an intervention; rather it needs to be planned for from the beginning to ensure processes are in place for ongoing inquiry, evaluation and learning.

The components for sustainability include:

- identification of the processes that were essential to maintaining and creating ongoing improvement

- embedding these processes in schools' "core business" as part of a coherent instructional programme

- having in place systems and processes to identify new challenges and how they will be acted on

- having a vehicle to systematically access and test knowledge that the cluster needs in order to continue improving outcomes.

In a following position paper, we will explain in greater detail (with examples) how to develop, monitor and evaluate a theory for improvement over time. Improvements appear to be sustained best when there is interdependence and coherence across different system levels, so we will also suggest possible roles and responsibilities for different levels of the system..

Professional learning communities

One possible vehicle for developing and enacting the components of the theories for improvement and sustainability is a professional learning community (PLC). These are groups of individuals who collaborate to advance their professional knowledge and skills to achieve valued outcomes. The purpose of the community is to learn from each other to achieve these outcomes. For example, the purpose of a community of researchers and schools could be to learn how best to implement a literacy intervention in schools to improve achievement.

Advancing professional knowledge and skills in a PLC can take many forms, depending on the purpose of the PLC. For example, a PLC set up to moderate writing assessments would focus on building a shared understanding of a quality production in order to ensure consistency in marking. On the other hand, a PLC set up to inquire into student data would focus on advancing the participants' knowledge and skills in diagnosing student needs from achievement data and relating those needs to classroom practices. In this sense, PLCs are vehicles to learn collaboratively and enact professional practices to achieve desired outcomes.

A PLC will only be useful if the members in the community learn from each other to advance their professional knowledge and practice. Learning in a PLC has two facets—the content of the PLC and the interpersonal interactions among the members of the PLC. The content is what is being learnt; for example, how to moderate writing. The interpersonal interactions are how members of the community learn the content from each other; for example, how the meeting is facilitated. The content for discussion can be strong and therefore useful to improve student outcomes, or weak, and not useful to improve student outcomes. The interpersonal interactions of the PLC can also be strong, with meetings organised in a way that members learn from each other; or weak, with meetings organised in ways that turn off members from learning from each other.

The optimal condition for learning is when both the content and the interpersonal interactions in the PLC are strong (see Table 2). If the content that is being discussed in the PLC is weak, then there will be no learning to improve valued outcomes, no matter how well the PLC works together. However, if the content is strong but the PLC does not function well, then learning will be limited.

Table 2 The relationship between content and interpersonal interactions

		Interpersonal interactions	
		Strong	Weak
Content	Strong	Optimal learning to improve valued outcomes	Minimal learning to improve valued outcomes
	Weak	Very limited learning to improve valued outcomes	Very limited learning to improve valued outcomes

In this section, we will not comment on the content being discussed in the PLC, as what we believe to be strong content is elaborated in other sections. For example, if the goal of the PLC is to learn how to inquire into data, then the content of that discussion should follow the inquiry cycle explained earlier.

Processes that are likely to encourage valuable learning in a PLC are:

- accessing and testing multiple sources of knowledge and skills

- collaborative inquiry including critical reflection on the ideas shared in the PLC

- developing shared understandings

- building collective efficacy

- building collective responsibility and collegial accountability.

Each of these points is now elaborated briefly. The PLC allows people of differing expertise—for example, expertise in analysing aggregated data, or expertise in teaching reading—to share their respective knowledge with the others. Evaluating the knowledge and skills shared in a PLC is fundamental to a PLC's effectiveness because not all knowledge and skills will contribute equally to the cluster's desired outcomes. So the PLC needs a way of testing the shared knowledge and skills, which involves collaborative inquiry into what is being shared through critical reflection. Such reflection includes testing the assumptions underpinning particular claims (e.g., Is it true that focusing on decoding is the ONLY way to raise reading achievement?) and checking that the knowledge or skills shared are more likely to result in the desired outcomes (e.g., How do we know that what person X is proposing is more likely to raise achievement than what we are currently doing?). This process of accessing and testing knowledge and skills helps create shared understandings in the PLC, thus enabling the PLC to create consistency and coherence across the community's practices.

The PLC further functions as the means for the community to develop collective efficacy - in other words, the collective belief that the community can achieve its desired outcomes. Similarly, the community can also be a source of building collective responsibility, in that members of the community take responsibility for each other's learning. Collective responsibility means that each member's goal is to try to advance the others' learning.

The position paper on professional learning communities will elaborate further on the processes of an effective learning community and provide transcripts of meetings to illustrate the qualities needed for these conversations. We will also suggest possible roles and responsibilities for different levels of the system, including schools and the MoE, to help optimise learning in a PLC through interdependence across these levels.

Planning to build evaluative capability

This fourth domain describes important elements of planning within schooling improvement. In essence this involves two processes. The first is to develop an effective and practical plan that assists participating schools to implement their agreed actions in the forthcoming cycle of activity. The second is for those involved to understand the principles underpinning effective planning processes so that, in the future, they are able to develop their own plans in association with, or independently of, external expertise. Learning how to plan by formulating a plan is likely to be effective only if the principles underpinning particular decisions are made explicit throughout the process. If plans are always developed and written by external experts, then the development of evaluative capability will be limited, with dependence rather than interdependence or independence developed.

One possible organising framework for planning is the cycles of inquiry illustrated in Figures 2 and 3. Schooling improvement is about learning how to do things better so the impact of schooling on student learning is greater than before. Such simple statements hide the complexity of this endeavour. A key message within the cycles of inquiry is that learning

needs to occur at all levels. Planning for schooling improvement, therefore, needs to recognise this complexity and identify these learning needs.

Like the principles underlying effective plans, there are some well-established principles of, and effective approaches to, schooling improvement. Making the knowledge underpinning these principles and approaches explicit through the planning process is also a way of developing the evaluative capability of those involved. The job of the school and/or cluster wanting to improve the percentage of students meeting or exceeding particular standards should *not* be to develop a plan independently of this knowledge, but to contextualise what is already known to ensure the plan meets local needs. For example, under certain conditions, explicitly evaluating teachers' existing instructional and curriculum knowledge and associated practices, and then engaging them in co-constructing their professional learning needs has proven to be effective across a number of interventions. Some of these conditions are:

> The job of the school and/or cluster is to contextualise what is already known to ensure the plan meets local needs.

- the knowledge and skills developed have proven effectiveness in raising student achievement

- the experts engaged have both high levels of pedagogical content knowledge and effective skills in facilitating professional learning

- the school culture and organisational arrangements promote professional learning (Timperley, 2008)

Reinventing wheels of intervention is typically ineffective and inefficient.

Ownership of plans is central to their successful implementation, and participation in the development of plans is one way to promote ownership. If schooling improvement efforts are to be successful, however, what is owned must be worth doing. Ownership must go hand in hand with knowledge of effective interventions if the plan is to achieve what it is designed to achieve. Ownership of the achievement problem, ownership of the data on current leadership and teaching practice, ownership of assumptions underpinning causal explanations and ownership of agreed actions are central to successful planning.

> It is important to ensure that what is owned is worth doing.

Ownership on its own, however, is insufficient to ensure any level of implementation, let alone effective implementation. Interdependence requires shared understandings about roles, responsibilities and accountabilities. If schools are working in clusters, for example, then cluster and school leaders need to clarify how the cluster plan can be translated into the school plans in ways that meet the school's learning priorities but remain coherent with the activities of other schools in the cluster. Part of the added value of a cluster is to have a

shared evidence base for student achievement and teaching and leadership practices so that the participating schools can consider what approaches have been most effective and what needs to change. Allowing schools to pick and choose from cluster initiatives without considering the need for coherence undermines these kinds of opportunities to learn from a shared evidence base.

Devolving responsibilities and accountabilities for implementation and monitoring of the plan to contracted co-ordinators can reduce the capability of schools and clusters. When formulating the plan, school leaders need to consider in advance what will happen when agreed actions are not implemented at either the school or classroom levels. Further information on suggested roles and responsibilities will be elaborated in the position paper on planning.

Māori Medium Education

Earlier we noted that in any schooling improvement approach there are features of particular significance to the context and identified Maori medium education (MME) as one such feature in New Zealand. MME as a generic term for a range of schooling provisions, including Kura Kaupapa Māori, Reo Rūmaki and Reorua. These options are also described as Kaupapa Māori education. An optimal model for schooling improvement in the context of New Zealand then is one that can enhance outcomes in

> An inclusive framework ensures that MME contexts, aspirations and valued outcomes are not forced to fit in as best they can.

MME. This requires the development of a theory for improvement that ensures that any improvement focus does not work in opposition to Māori language and cultural development, which is fundamental to MME contexts. Effective schooling improvement also ensures that "achievement" is inclusive of the kaupapa of MME and Kura Kaupapa Māori. A model that is inclusive makes genuine space for kaupapa and aspirations contained within, for example, Te Aho Matua and Ka Hikitia (Ministry of Education, 2008). Each layer of the education system is responsible for ensuring that this happens.

Difficulties identified in Māori medium settings that participate in schooling improvement include relatively limited availability of appropriate assessments that capture the kaupapa of MME and kura kaupapa Māori, and limited teacher knowledge about using available assessments, needed in order to ensure a strong evidence base. Schooling improvement provides potential opportunities to do at least two things. One is to identify if tools need to be developed. The current work on developing national standards also provides another opportunity for this. Another is to support the professional capacity building required to develop and to use tools effectively to improve teaching and learning. The MOE (and its contracted developers and suppliers of assessment tools), have a particular role in and responsibility for ensuring that the development and the use of assessment tools does not

arbitrarily and uncritically assume MME pedagogical and cultural compatibility with English medium education (Rau, 2001).

MME contexts can participate in schooling improvement projects in various arrangements. Some involve kaupapa clustering with other MME schooling

> Identifying and examining the potential opportunities and potential costs of a clustering arrangement is of particular importance with regard to MME schooling improvement.

provisions, others involve mainly geographical clustering with English medium schooling. The various arrangements carry potential opportunities and potential costs for interdependence and for capability. These need to be identified and examined carefully, particularly at school and cluster level.

Our theory of learning from the position papers

The model for schooling improvement outlined in this paper is intended to provide a framework for other papers. Pieces of paper on their own, however, have little influence as mechanisms for change unless other conditions are met.

One condition is that the papers undergo critique and their meaning is unpacked through discussion involving the evaluation team, the MOE leaders and members of the clusters. No doubt revisions will follow as a result of such discussion that will make the papers more useful to the current schooling improvement clusters.

> Pieces of paper are poor mechanisms for change. They must be accompanied by other opportunities to learn.

A second condition is that those responsible for leading schooling improvement throughout the country (such as MOE and contracted co-ordinators) develop sufficient understanding of the revised papers to enable them to work with their clusters in the ways described. The investment in professional learning to develop this level of understanding is likely to be considerable if more than superficial implementation is to be achieved. Typically, deep understanding develops through engagement in ongoing cycles of development, inquiry and implementation as shown in Figures 2 and 3.

A third condition fundamental to success is that revised versions of the papers become part of the routines involved in planning and reviewing schooling improvement work at the individual schools and clusters. At the same time, explicit routines need to be developed to ensure effective practices are embedded within all schooling improvement activities. Unless the papers become an integral part of schooling improvement work, they are likely to be read and shelved, with business as usual continuing.

References

Alton-Lee, A. (2003). *Quality teaching for diverse students in schooling: Best evidence synthesis*. Wellington: Ministry of Education. Available at: http://www.minedu.govt.nz/index.cfm?layout=document&documentid=8646&data=1

Bandura, A. (1995). Exercise of personal and collective self efficacy in changing societies. In A. Bandura (Ed.), *Self efficacy in changing societies* (pp. 1–45). Cambridge: Cambridge University Press.

Bryk, A., Sebring, P., Allensworth, E, Luppescu, S., & Easton, J. (2007). *Organizing schools for improvement.* Chicago: University of Chicago Press.

Katz, S., Earl, L., & Ben Jafaar, S. (in press). *Building and connecting learning communities: The power of networks in school improvement.* Corwin Press.

Lai, M. K., McNaughton, S., Timperley, H., & Hsiao, S. (2009). Sustaining continued acceleration in reading comprehension achievement following an intervention. *Educational Assessment, Evaluation and Accountability.* Retrieved 19 February 2009, from http://www.springerlink.com/content/m056j60853pq69h1

McNaughton, S., & Lai, M. K. (2009). A model of school change for culturally and linguistically Diverse students in New Zealand: A summary and evidence from systematic replication. *Teaching Education, 20*(1), 1–21.

Ministry of Education. (2008). *Ka hikitia—managing for success: The Māori Education Strategy 2008– 2012*. Wellington: Ministry of Education.

Rau, C. (2001, September). *He Ara Angitu: A framework for capturing the literacy achievement of year one tauira in Māori medium.* Paper presented at Assessment Hui for the Ministry of Māori Education, Wellington.

Robinson, V., Hohepa, M., & Lloyd, C. (in press). *School leadership and student outcomes: Identifying what works and why. A best evidence synthesis*. Wellington: Ministry of Education.

Robinson, V., & Lai, M. K. (2006). *Practitioner research for educators: A guide to improving classrooms and schools.* Thousand Oaks, CA: Corwin Press.

The *International Handbook of School Effectiveness and School Improvement*, edited by T. Townsend has a whole section on leadership with articles by Moos and Huber (pp. 579–596); Marzono (pp. 597–614); Leithwood (pp. 615–634); Silins and Mulford (pp. 635–658); and Walker, Hallinger, and Qian (pp. 659–680).

Timperley, H. (2008). *Professional learning and development*. International Bureau of Education / International Academy of Education. Available at: http://www.ibe.unesco.org

Timperley, H., Parr, J., Hohepa, M. LeFevre, D., Lai, M., Dingle, R., Schagen, S. (2008). *Findings from an Inventory Phase of the Building Evaluative Capability in Schooling Improvement Project.* Report to the Ministry of Education, Wellington, New Zealand.