



School quality in Australia

Exploring the drivers of student outcomes and the links to practice and school quality

FINAL REPORT – December 2017
Department of Education and Training

© Commonwealth of Australia 2017

With the exception of Deloitte branding, content provided by third parties, and any material protected by a trademark, all textual material presented in this publication is provided under a Creative Commons Attribution 4.0 International licence (CC BY 4.0) <creativecommons.org/licenses/by/4.0/>. You may copy, distribute and build upon this work for commercial and non-commercial purposes; however, you must attribute the Commonwealth of Australia as the copyright holder of the work. Content that is copyrighted by a third party is subject to the licencing arrangements of the original owner.



This report was commissioned by the Australian Government Department of Education and Training. The findings and views expressed in this report are those of the authors and do not reflect the views of the Department of Education and Training.

Suggested citation:

Deloitte Access Economics (2017). *School quality in Australia: Exploring the drivers of student outcomes and the links to practice and school quality*, Canberra: Australian Government Department of Education and Training.

ISBN 978-1-76051-352-8 [PDF]

ISBN 978-1-76051-353-5 [DOCX]

Liability limited by a scheme approved under Professional Standards Legislation.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity.

Please see www.deloitte.com/au/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms.

Contents

Definitions	i
Glossary	iii
Executive summary	i
1 Introduction	1
1.2 A measurement framework for drivers of student outcomes and school quality	2
1.3 The structure of this report	4
2 The drivers of school quality – establishing a measurement framework	5
2.1 The factors which contribute to student outcomes – a brief literature review	8
2.1.1 Contributions from the student and home	8
2.1.2 Contributions from the classroom and the teacher	11
2.1.3 System level factors	16
2.1.4 Implications for this study’s methodological approach	20
2.2 How school performance and school quality is measured	21
2.3 A structural equation to isolate the drivers of school quality	23
3 Empirical methodology and data	24
3.1 The PISA and TIMSS datasets	24
3.2 Understanding and measuring school quality	26
3.3 Identifying and measuring the drivers of school and classroom quality	27
3.3.1 Mapping questions to themes and clustering	28
3.3.2 Identifying key questions for modelling and analysis	33
3.4 Providing a link between system settings and jurisdictional performance	34
3.4.1 Revealing systemic changes in performance	34
3.4.2 Desktop review of schooling system settings and initiatives	35
3.5 Limitations to the analysis	36
4 Findings from the analysis	38
4.1 Measuring school quality and value-added	38
4.2 Identifying and measuring the drivers of school quality	44
4.3 Drivers of school quality and their relative importance	45
4.4 Relative importance of school quality drivers across outcome measures	53
4.5 Relative importance of school quality drivers across geographies	54
5 Implications for government and future research directions	56
5.1 An evidence base on schooling systems, practice and performance in Australia	56
5.1.1 Results of empirical analysis of drivers of school quality	56
5.1.1 A framework for understanding the role of government in the school system	57
5.1.2 The current state of schooling policy in Australia	60

5.2	The importance of evaluation	61
5.2.1	Possible implications for policy-makers	62
5.3	Conclusions and future research directions	64
	References	66
	Appendix A : Detailed measurement frameworks, for mapping of PISA and TIMSS questions to themes	73
	Appendix B : Evidence for Learning (<i>Teaching and Learning Toolkit</i>)	80
	Appendix C : Detailed modelling results	82
C.1.	Controls	82
C.2.	Value added results	82
C.3.	Drivers of school quality (PISA)	85
C.4.	Geographic analysis	111
	Appendix D : Providing a link between schooling system settings and performance across Australian jurisdictions	114
D.1.	Trends in performance across Australian jurisdictions	114
D.2.	Identification of key system settings	115
	Limitation of our work	132
	General use restriction	132

Definitions

Term	Definition
Programme for International Student Assessment (PISA)	PISA is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students. Three domains are tested: reading literacy, mathematical literacy and scientific literacy. Alongside the test, students, school leaders and teachers complete surveys about aspects of their school and practice.
Trends in International Mathematics and Science Study (TIMSS)	TIMSS (Trends in International Mathematics and Science Study) is a large-scale assessment designed to inform educational policy and practice by providing an international perspective on teaching and learning in mathematics and science. First conducted in 1995, TIMSS reports every four years on the achievement of Year 4 and Year 8 students. Background questionnaires are also conducted with students, teachers and school leaders.
School quality	<p>School quality is defined as the contribution that a given school makes to the outcomes of its students, after controlling for contextual characteristics. The quality of a school is therefore a driver of student outcomes, as are other factors relating to the individual characteristics of students, or the context of the school and/or education system.</p> <p>Drivers of school quality are defined as the attributes of a school's practice and management that drive student outcomes, and therefore determine a school's 'quality'. These attributes of practice and management are defined as factors over which a school has some degree of control—in contrast to a school's contextual characteristics.</p>
Classroom quality	Classroom quality is defined as the contribution that a given teacher and classroom structure/environment makes to the outcomes of its students, after controlling for contextual characteristics. The quality of a school is therefore a driver of student outcomes, as are other factors relating to the individual characteristics of students, or the context of the school and/or education system.
Prior Academic Achievement and self-efficacy	Prior Academic Achievement and self-efficacy describes the level of prior achievement a student brings to the classroom and the intrinsic measures of engagement and wellbeing associated with learning in the classroom.
Teaching Efficacy	Teaching efficacy is an all-encompassing term that captures the quality of a teacher and the quality of their teaching. In particular, it captures teacher attributes (such as qualifications) and teaching practice—which includes teachers' approaches to organisation of teaching and learning in the classroom.
School leadership, governance and culture	School leadership, governance and culture relates to school's approaches to managing, governing and leading schools, and the culture of leadership that is developed. This may include the nature of school's models of distributed leadership, as well as the approaches of school principals and leading teachers towards school wide instructional approaches, and approaches to student management.

Term	Definition
Classroom organisation and environment	Classroom organisation and environment covers how different aspects of classroom conditions and dynamics (context, organisation and, environment) influence student outcomes.
Autonomy	Autonomy is used to describe the level of local decision-making authority provided to schools within a given education system (including aspects of school management and administrative decisions, recruitment of teachers, and curriculum).
Accountability	Accountability relates to how and to what extent schools are held accountable for student outcomes, and for the quality of their practice and management more broadly.
Resourcing	Resourcing relates to the type and level of resources schools may use to deliver teaching and learning activities (and other activities related to the running of a school).
Student outcomes	Include any positive outcome for students associated with schooling, which may include subsequent labour market outcomes or other quality of life measures. In this report the focus is on measures of student academic achievement provides test score results.
Intermediate outcomes	Intermediate outcomes are outcomes which are driven by factors that affect school performance but are also drivers of school performance in and of themselves. For example, students are more likely to be engaged in school if they have teachers using effective practices; this engagement drives their performance in overall school outcomes measures.

Glossary

Acronym	Definition
ACARA	Australian Curriculum, Assessment and Reporting Authority
ACT	Australian Capital Territory
ESCS	Economic Social and Cultural Status
NSW	New South Wales
NT	Northern Territory
OECD	Organisation for Economic Co-operation and Development
PISA	Programme for International Student Assessment
QLD	Queensland
SA	South Australia
SES	Socio-Economic Status
TAS	Tasmania
TIMSS	Trends in International Mathematics and Science Study
VIC	Victoria
WA	Western Australia

Developing an evidence base linking policy, practice and performance in Australian schools

1. The challenge and opportunity

Australia is falling behind on international tests of student achievement even though school funding is growing.

We have been monitoring outcomes nationally for nearly a decade



... but without a **consistent evidence base** to inform **strategic focus** on areas where changes to **practice** can make a difference to outcomes, there is limited assurance that these investments will translate into improvements in school quality.

State and Commonwealth governments have implemented **hundreds of initiatives** and **significantly increased funding** over the last 20 years



There is an opportunity to improve—if we were able to **lift the performance of a school from the bottom 10% to the top 10% of schools**, it would be equivalent to adding an additional **1.5 years** of learning for those students.

And by increasing Australian student outcomes by just 5% nationally, we could:



Grow Australia's economy by \$26bn

2. A new measurement framework

This research provides guidance for policy by identifying the aspects of classrooms and schools that matter most to student outcomes.

By mapping indicators of **context, practice, policy, and outcomes** from the PISA and TIMSS datasets into a robust and consistent framework, this study brings together the **most recent Australian data on the drivers of school quality**.

School quality is defined as **the impact that school practice has on student outcomes**. Observed drivers of school quality include:



School resourcing
School autonomy
Teaching efficacy

Classroom organisation and environment
School leadership

This framework and methodology allows for the empirical measurement of the relative and absolute importance of these drivers in **explaining variations in outcomes** across Australian schools and classrooms. These outcomes include:



Academic achievement
Engagement
Sense of school belonging
Safety and wellbeing

3. A new evidence base to inform policy

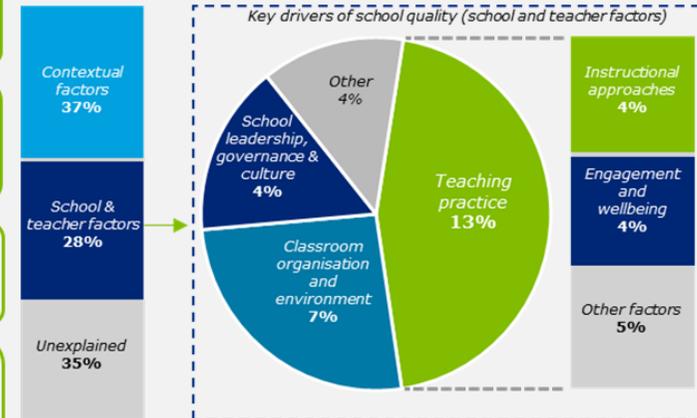
School and teacher factors contribute as much as **28%** of variation in students outcomes.

Within this 28%, **teaching practice, classroom organisation and school leadership** are the most important drivers.

The most important drivers within **teaching practice** are **instructional approaches, and engagement and wellbeing**.

These results are broadly consistent across outcome measures, however **school leadership** and **school resourcing** appear more important when explaining student engagement and wellbeing outcomes.

Contribution to variation in TIMSS Year 8 Maths scores, 2015



While this evidence confirms what many already know about importance of practice in the classroom, it is now possible to demonstrate that teaching practice is at least **twice as important** in explaining student outcomes when compared to all other observed drivers of school quality.

4. Implications and future directions

This work illustrates the **potential gains** to be made from Australian schooling policies that focus on improving the **most important drivers of school quality**—teaching practice, classroom organisation & environment, and school leadership.

These gains will only be achieved by focusing the policy levers of government on the factors which matter most—those which influence **practice in the classroom**. However, these critical drivers are often the most removed from system level policy settings and levers. This means Australian schooling needs **better evidence that links practice at the school and classroom level to policy and performance**.



With this evidence base, government may be supported in...

1. Demonstrating the link to improvements in teaching practice before investing in initiatives intended to drive improvements in outcomes
2. Setting a strategic and long-term focus on the outcomes impact of sustained changes to practice at the classroom level
3. Relying on, and encouraging the use of, evidence-based interventions across Australian classrooms and
4. Collecting and sharing data and evidence on how interventions result in improved practice.

This study provides an evidence-based framework to support government in understanding how policy and reform may be aligned with evidence of what matters to improving student outcomes. This evidence may be used to develop a formative framework for improving student outcomes to underpin policy going forward—including to help drive more effective uses of funding in Australian schools.

Executive summary

Australia's schooling system has consistently been ranked as one of the highest performing in the world. This performance is underpinned by the quality of schools, the quality of teachers and the effectiveness of Australian teaching practice. However, there is scope for Australia's performance to improve. Most recently, the 2015 Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) tests have indicated Australia's schooling results declining in both absolute and relative terms.

Reversing these trends will require governments to direct resources and policy efforts toward the endeavours of school practice most capable of lifting quality and student outcomes. However, there is currently limited empirical evidence about how policies, schooling practices and student factors are associated with these school quality trends in the Australian context and what this means for Australia in lifting school quality.

There is therefore a need to better understand the underlying drivers of student outcomes in Australian schools—particularly as they relate to practice and policy—in order for government investments and policy to be better targeted to drive improvements in the performance of Australia's schooling system.

The Department of Education and Training ('the Department') engaged Deloitte Access Economics to undertake research into the drivers of school quality in Australia, with the goal of expanding the evidence base available to the Department in analysing, evaluating and developing school policy. In particular, this work provides a new empirical framework for assessing these drivers in the Australian context, and presents the results of this framework using PISA and TIMSS data.

Establishing a new evidence base to inform policy-making

Studies such as Hattie's *Visible Learning (2009)* have provided a comprehensive overview of the factors that influence student outcomes in school education, including the relative effects of different contextual attributes, policies and practices on student learning. These studies draw on extensive international and Australian research to establish a detailed evidence base capturing what works and, equally importantly, what does not work, in school education.

However, there remains a gap in the evidence base to support findings on the drivers of school quality in Australia. This arises from the fact that:

- While this existing evidence base has helped us understand the factors that drive educational outcomes in schools, the link between this evidence base and the role of government is often less clear and few researchers seek to provide a summative structure which relates their findings to implications for government.
- Further, this evidence typically relies on a range of studies conducted on students in different educational contexts and systems which make meaningful comparisons of their relative effects on a consistent measure of student outcomes hard to establish.

Few studies have sought to conduct empirical analysis of the drivers of student outcomes, and school quality, using a comprehensive dataset of Australian students and schools. Similarly, few studies seek to use such an evidence base to link the drivers of school quality to the role of government.

This research study is differentiated from previous research as it seeks to use a consistent set of Australia-specific evidence on student outcomes and school practice to provide insights on the drivers of school quality within a single consistent framework. This makes it possible to identify the relative effect sizes of different drivers of school quality, information that is not presented by the existing literature.

Any analysis of this nature is limited by the quality of the data available. This study utilises indicators of practice contained in the PISA and TIMSS datasets to match particular aspects of

practice to observed student outcomes. The accuracy of the results is therefore limited in part by the quality of the matching exercise facilitated by these datasets and the quality and appropriateness of the questions that comprise these datasets. Hence, this report provides evidence on the most important drivers of school quality, while noting that the approach may be further refined over time as better evidence becomes available.

Providing a link between practice and performance

The goal of this study is to develop a framework that links elements of practice and student outcomes, and which is capable of utilising existing data sources to empirically estimate this link, in order to provide evidence of:

- The factors that drive student outcomes in Australia, and in particular the role that school quality plays in driving student outcomes;
- The attributes of practice that define school quality, and how these attributes may be measured; and
- The role that governments can play in influencing school practice and driving improvements in school quality.

It is possible to organise the contributing factors to student learning outcomes into three broad categories:

1. **Student factors**—including factors that relate to a student’s background, and context, as well as measures of prior achievement and self-efficacy in learning.
2. **School level factors**—including aspects of teaching efficacy, school practice and management, and other aspects of school practice.
3. **System level factors**—including characteristics of schooling systems, such as autonomy, accountability and resourcing.

Within each of these categories is a range of factors that are known to influence student outcomes. Some of these (such as socio-economic status) are identified and measured in international datasets such as PISA and TIMSS.¹ Others may be identified through external evidence sources, such as system level attributes relating to policy and regulatory settings.

A simple overview of the approach

The most robust way to determine the relative effectiveness of different drivers of school quality is to regress a measure of student outcomes (such as test scores) on these drivers, while controlling for other contextual factors. Unfortunately it is difficult to observe each driver for an individual student, meaning that this simple regression analysis cannot be undertaken.

However, the PISA and TIMSS datasets do contain other indicators that are observed for each student. By mapping these indicators to the (unobserved) drivers of quality, it is possible to use them in the regression analysis in place of those drivers. The core contribution of this work is to undertake this mapping in a way which allows the derivation of empirical estimates of the relative importance of the different drivers using a robust and consistent framework.

The methodology for this study involves providing a link between practice and student outcomes by bringing together evidence from literature and previous research, and original empirical analysis of the PISA and TIMSS datasets. This is made possible through the inclusion of questions in each of these datasets that provide indications of certain aspects of practice, thereby providing a measurable link between practice and student outcomes. More specifically, the approach involves the following steps:

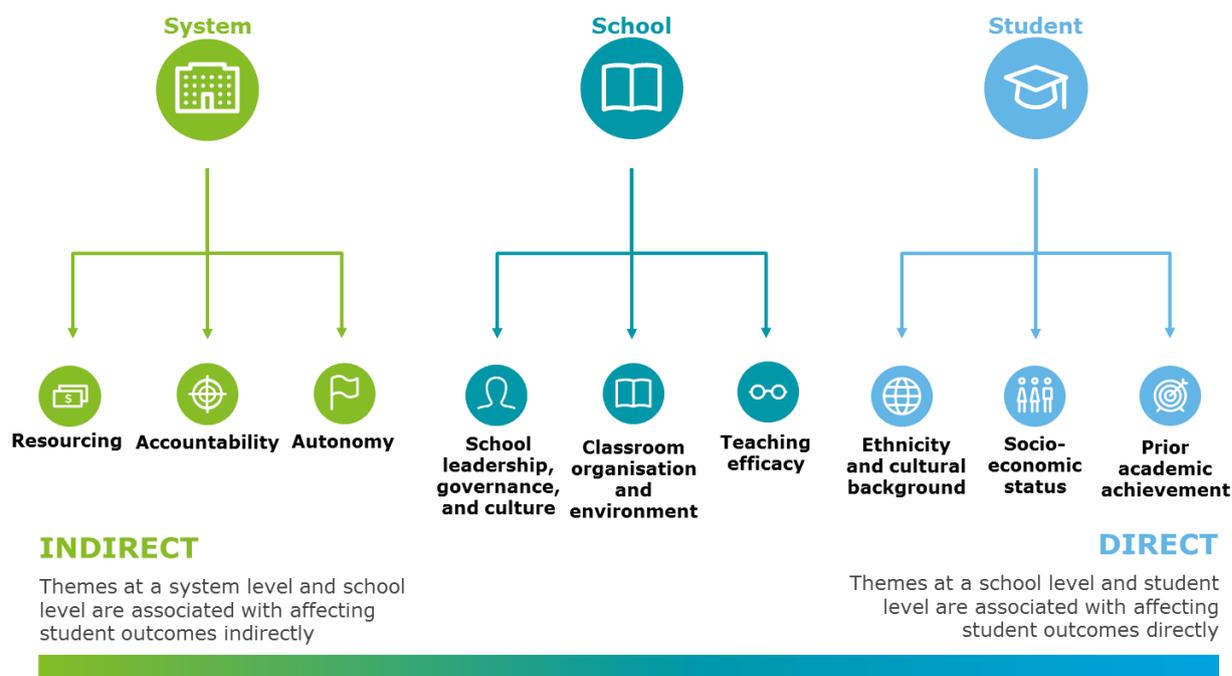
- A literature review is undertaken to identify the key drivers of student outcomes in existing research. These drivers are then categorised into themes and sub-themes.
- The questions contained in the PISA and TIMSS datasets are then mapped to these themes, effectively creating proxy variables in these datasets for each of the themes.

¹ It should be noted that PISA and TIMSS use different measures of these variables. PISA constructs its own proxy index for measuring socio-economic status (known as the ESCS Index). TIMSS collects information through a number of proxy variables, including the number of books in a student’s household.

- Student outcomes are then regressed against these drivers to empirically determine which elements of practice have the greatest impact on student outcomes. These identified aspects of practice are defined as the *drivers of school quality*.

Figure i below provides a high-level overview of this measurement framework, which relates system, school, and student level factors to a set of nine *anchor themes*. These themes represent the key drivers of student outcomes that have been identified through a targeted review of the leading literature on what matters in school education internationally.

Figure i: Overarching measurement framework for the empirical analysis



Source: Deloitte Access Economics

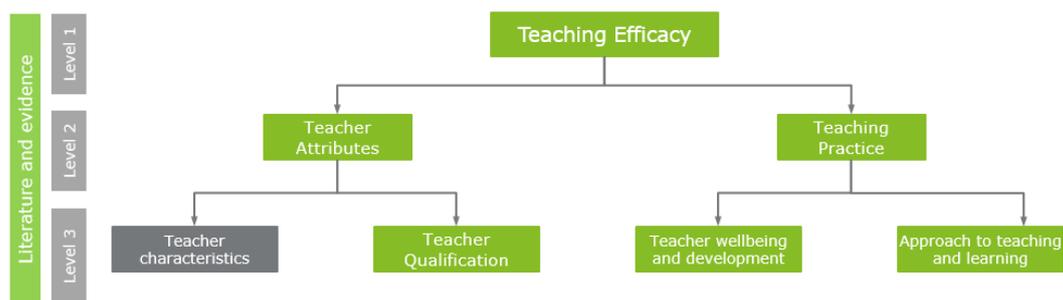
Key literature review findings

Evidence from the literature clearly shows that teaching efficacy is among the most influential factors driving educational outcomes, while school leadership, governance and culture at a school level generally influence student outcomes through their impact on the classroom learning environment and the quality of teaching. The factors that influence teaching efficacy are complex and multi-faceted. Broadly, the literature indicates that teaching efficacy can be conceptualised broadly by two themes—**teacher attributes** and **teaching practice**.

For the purposes of categorising questions and measures within the PISA and TIMSS datasets, this framework is subsequently explored in greater levels of granularity. For example, the theme of *teaching efficacy* (which is broadly synonymous with the notion of 'teaching quality') is separated into *teacher attributes* and *teaching practice*. These separate themes distinguish between the skills and qualifications of teachers which inform the effectiveness of their teaching practice; and their teaching practice itself, which includes factors such as ongoing professional development and approaches to teaching and learning in the classroom—as outlined in Figure ii below.

In broad terms, these different factors or *themes* can be classified as being: **contextual**, in the sense that they relate to the attributes that students bring with them to the classroom or other contextual circumstances relating to the school environment; or factors which relate more to the **quality of schools** and the **performance of systems**, which capture elements of educational practice that can be influenced by government and can be connected to the quality of teaching and learning in the classroom.

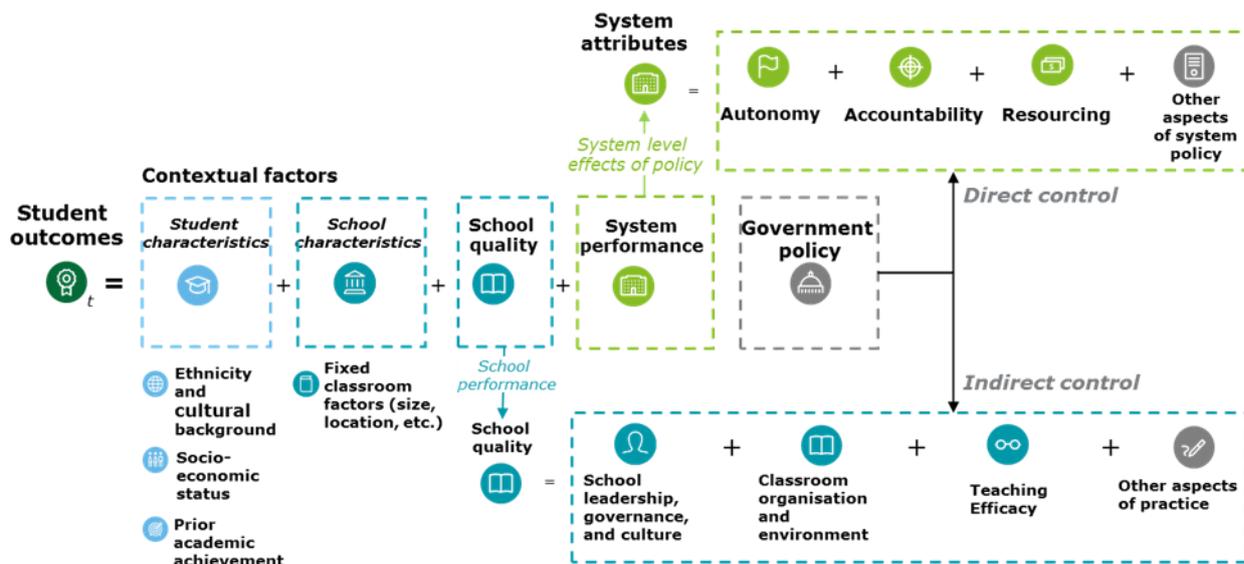
Figure ii: The components of teaching efficacy



Source: Deloitte Access Economics

By separating the drivers in this way it is possible to isolate the variation in outcomes explained by individual school practice, separate from their contextual environment (for example, the socio-economic status of their local community), the system (for example, government or non-government) or jurisdiction (that is, state or territory) in which they operate and seek to identify the observable characteristics that explain these variations. A conceptual illustration of this empirical approach is provided in Figure iii below.

Figure iii: Conceptual illustration of empirical approach



Source: Deloitte Access Economics

As with any original empirical study, the approach used in this report has its limitations. In particular:

- While the PISA and TIMSS tests and surveys are conducted every 3-4 years, they are not longitudinal in nature. That is, the students and schools sampled for the test are not common across years. This means that it is not possible to capture dynamic effects of practice on student outcomes over time and that the analysis can only be conducted with contemporaneous observations of school practice and student performance.
- Following from the static nature of the PISA and TIMSS tests, it is not generally possible to account for students' prior achievement when measuring the effect of practice on outcomes. This may overstate the effects of aspects of school quality on student outcomes relative to their actual effect, particularly in the presence of ability-based school selection policies.

- The strength of the conclusions made in this report rely on the quality of the evidence available in the PISA and TIMSS datasets and the way in which the data is used to represent the measurement framework developed through this study. While these datasets are extensive, they are by no means comprehensive. In some cases, where the evidence shows that a given theme is more or less important in explaining student outcomes, this may result from the fact that no effective instrument was available to demonstrate its impact.
- The PISA and TIMSS tests each have different scope: PISA is conducted with 15-year old students across several year levels, and tests the ability of students to solve real-world problems in English, Mathematics and Science. TIMSS, on the other hand, is aimed at Year 4 and Year 8 students and is based on school curriculum in those year levels. Comparisons of the results of the analysis *between* TIMSS and PISA scores should be made in light of this.

The contribution of school quality in Australia

Within the context of available data, it is possible to measure the contribution of school quality to student performance by estimating the variation in student outcomes that remains after controlling for the observable contextual characteristics of students and the school (recognising that it is not possible to perfectly control for all factors that may influence student outcomes—in particular, students’ innate ability or prior learning outcomes). That is, after controlling for other contextual factors, what remains unexplained can be interpreted as the effect of various individual school factors. This is achieved through the use of value-added models (also known as multi-level models) that isolate the effects of differences in school practice on student performance, while controlling for observable characteristics (such as socio-economic status).

By controlling for observed student and school characteristics, comparisons of performance between schools are made on a ‘like-for-like’ basis—which is equivalent to, for instance, measuring the effect on a given student’s learning outcomes of moving that student from one school to another. In this sense, it is a measure of the school’s attributable impact on learning outcomes—also known as ‘value-added’.

The relative importance of drivers of school quality

PISA and TIMSS also incorporate additional questionnaires to gain insights into the contextual factors associated with student achievement. In PISA, students and school leaders are surveyed; in TIMSS, students, school leaders and teachers are surveyed. PISA contains over 400 individual questions related to teaching practice and school conditions, while TIMSS contains over 250 questions.

Mapping questions to themes

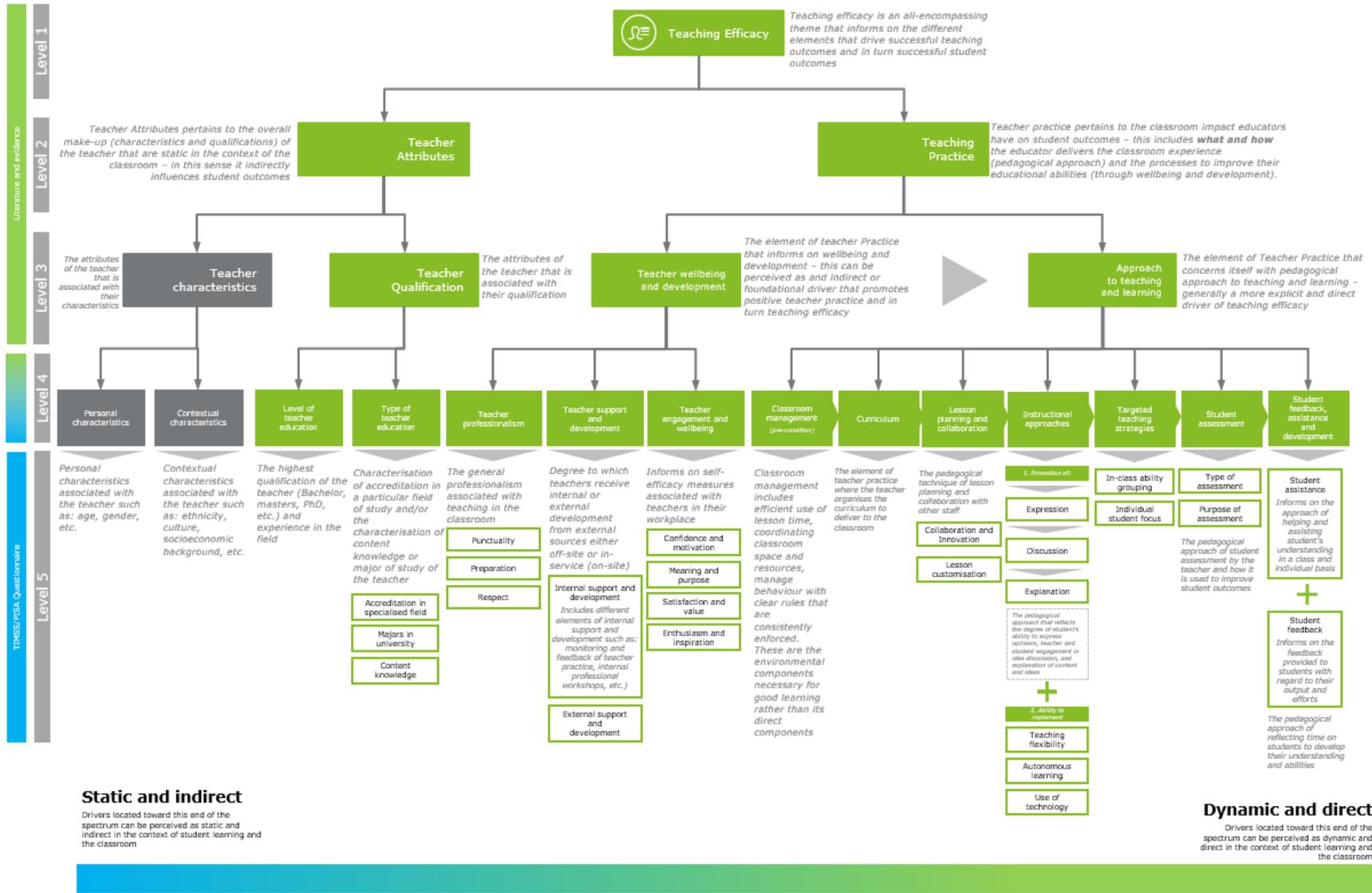
As described earlier, key PISA and TIMSS questions are mapped to analytical themes identified through the literature review. This is achieved by analysing all of the questions in each of these datasets and organising them into subordinate themes with the conceptual structure established above.

An illustration of the outcomes of this mapping process is presented in Figure iv for the theme of *Teaching Efficacy*. Each of the boxes in this diagram represents a thematic ‘driver’ which comprises a group of PISA and TIMSS questions capturing the same general theme. These variables, constructed through the mapping of the PISA and TIMSS questions to the measurement framework, form the basis of the empirical modelling.

Clustering and selecting representative questions for analysis

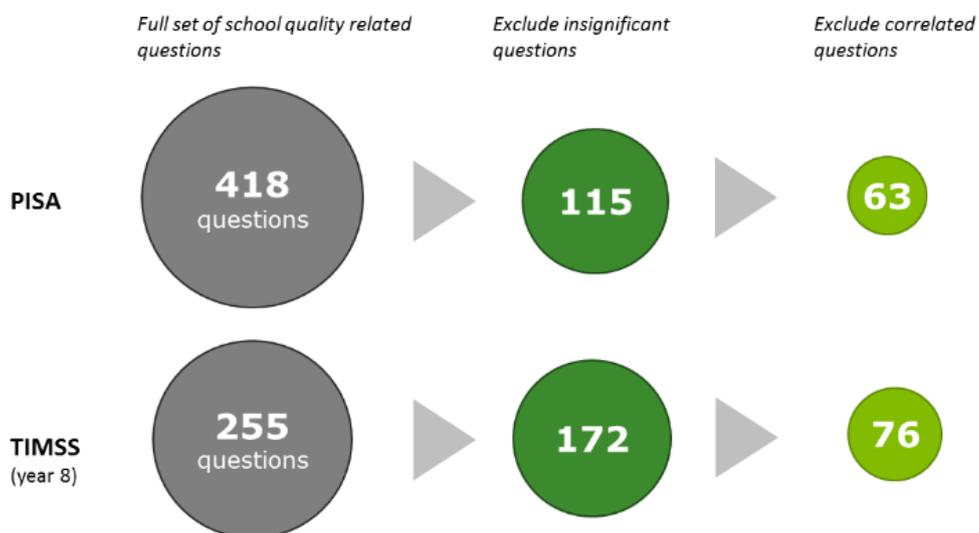
Making the modelling approach tractable requires reducing the number of questions used in the analysis. This is achieved by excluding those questions that are found to be insignificant in predicting variations in student outcomes, and—within each conceptual theme—excluding questions that are highly correlated with the most significant ‘representative question’ in that theme. Overall, the 418 questions of the PISA dataset are reduced to 63 representative questions that represent 22 defined sub-themes relating to quality at the school and system level. A similar process is undertaken for TIMSS, reducing the set of 255 questions to 76 questions representing 24 sub-themes, as outlined in Figure v below.

Figure iv: Measurement framework mapping of Teaching Efficacy theme from literature and survey questions



Source: Deloitte Access Economics

Figure v: Flow chart for selecting representative questions



Source: Deloitte Access Economics analysis of PISA and TIMSS data

This set of representative questions is then added to a regression model of student outcomes, alongside the established set of contextual controls. This regression model is used to understand the relative importance of each driver of school quality in predicting student outcomes, and to reveal the key drivers of school quality in Australian schools.

Findings from the analysis

The estimated drivers of school quality can be ranked by their ability to explain variation in student PISA or TIMSS scores. Drivers that explain more of the variation in scores are necessarily those that have the biggest impact on scores.

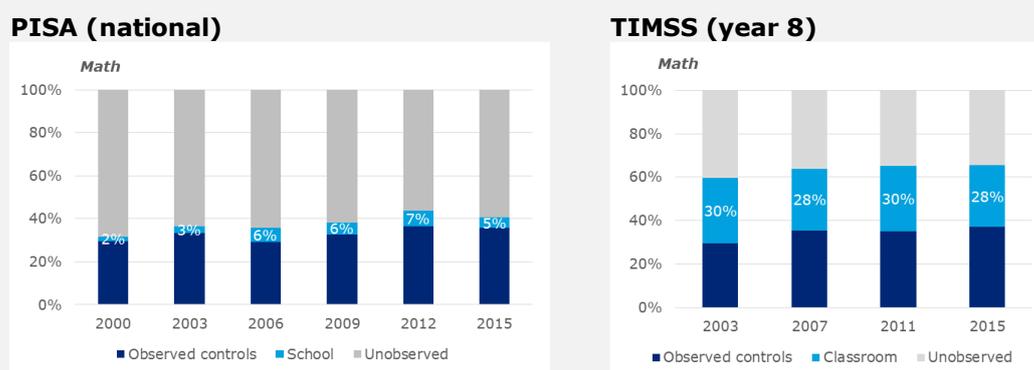
Key empirical findings

The results from the analysis of the PISA dataset demonstrate that **variations in 'school quality' in Australia explain between 2% and 7% of the total variation in student outcomes observed across the country**. This is in comparison to observable student and school level contextual characteristics (such as socio-economic status), which explain around 30% of the variation in student outcomes, as outlined in Chart i below.

Using the TIMSS data, the analysis can focus on variation between individual classrooms (including their teachers), rather than variation between schools. As outlined in Chart i below, the contribution made by **classroom quality to student outcomes is typically higher than the contribution made by school quality**. This finding is consistent with evidence from similar studies which emphasise the significance of individual teaching practice in driving student outcomes, irrespective of the specific school environment.

It should also be noted that PISA and TIMSS are different assessments in terms of the skills examined. In particular, noting that TIMSS is a curriculum based test (while PISA measures student skills in reading, maths and science) and that the age and grade of students is higher in PISA than TIMSS it may be expected that classroom related factors (such as teaching practice) would be more important in TIMSS than PISA.

Estimates of the contribution of school quality, classroom quality and other factors to student outcomes



Source: Deloitte Access Economics analysis of PISA and TIMSS data

Table i below lists the relative importance of nine high level themes across the PISA and TIMSS datasets. The results in this table and can be interpreted as the proportion of variation in outcomes explained by each of the themes. For instance, variation in *teaching practice* explains the greatest variation in student scores, at 6.1% for PISA maths scores, and up to 13.1% for TIMSS math scores.

Table i: Relative importance of drivers of school quality themes (2015)

Themes	PISA (maths)	TIMSS (maths, year 8)	TIMSS (maths, year 4)
All school drivers	13.8%	27.5%	8.4%
Teaching efficacy - Practice	6.1%	13.1%	3.9%
Classroom organisation and environment - Environment	4.8%	3.9%	0.9%
School leadership, governance and culture	2.1%	4.5%	2.0%
Resourcing - Material based	0.5%	2.2%	0.7%
Autonomy	0.4%	n/a	n/a
Classroom organisation and environment - Organisation	0.2%	3.6%	1.1%
Teaching efficacy - Attributes	0.1%	1.4%	0.3%
Accountability	0.0%	n/a	n/a
Resourcing - Curriculum and staff based	n/a	0.3%	0.2%

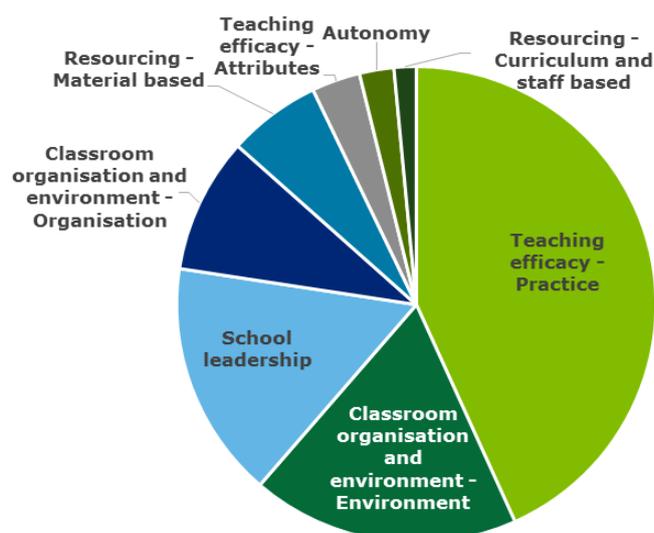
Source: Deloitte Access Economics analysis of PISA and TIMSS data

This result is not unexpected, based on findings from the literature about what matters in schooling education. However, the magnitude of the importance of *teaching practice* relative to other factors, such as *school leadership, governance and autonomy*, is notable. Indeed, the most important drivers of student outcomes and school quality, ranked in order, are:

- Teaching efficacy - practice
- Classroom organisation and environment - Environment
- School leadership, governance and culture
- Classroom organisation and environment - Organisation
- Resourcing - Material based resourcing

An illustration of the relative importance of these drivers is provided in Figure vi below. This pie-chart shows the share that the school quality 'themes' contribute to the overall contribution made by school quality drivers (that is, the relative proportion of the variation in outcomes explained by each driver), averaged across the PISA and TIMSS datasets.

Figure vi: Illustration of the overall relative importance of quality drivers (average across PISA and TIMSS)



Source: Deloitte Access Economics analysis of PISA and TIMSS data

The relative importance of the sub-themes within *teaching practice* is presented in Table ii below. *Instructional approaches* describe the pedagogical approaches utilised by teachers which relates to the degree of students' ability to express opinions, teacher and student engagement in idea discussion, and explanation of content and ideas. This sub-theme is found to have the largest explanatory power for high school students in PISA and TIMSS. For primary schools, variations in *instructional approaches* are less important drivers of outcomes than teaching practice relating to *curriculum* and aspects of *teacher engagement and wellbeing*.

Table ii: Relative importance of themes – teaching practice (2015)

Sub-themes (Teaching practice)	PISA (maths)	TIMSS (maths, year 8)	TIMSS (maths, year 4)
Approach to teaching and learning - Assessment	0.09%	1.94%	0.02%
Approach to teaching and learning - Instructional approaches	2.85%	4.28%	0.92%
Approach to teaching and learning - Curriculum	n/a	2.83%	1.18%
Approach to teaching and learning - Lesson planning and collaboration	n/a	1.03%	0.25%
Approach to teaching and learning - Targeted teaching strategies	2.00%	n/a	n/a
Wellbeing and development - Professionalism	3.06%	0.20%	0.46%
Wellbeing and development - Engagement and wellbeing	n/a	4.02%	1.18%
Wellbeing and development - Support and development	0.08%	0.16%	0.07%
Total teaching efficacy - practice	6.07%	13.11%	3.88%

Source: Deloitte Access Economics analysis of PISA and TIMSS data

Overall, these results emphasise the importance of key aspects of teaching practice, such as targeted teaching and effective instructional approaches. Notably, these pedagogically focused themes are more important than those themes which relate to curriculum and assessment, and the process of lesson planning and collaboration.

Implications for the role of government in Australia

This work illustrates the **potential gains** to be made from Australian schooling policies that focus on improving **the most important drivers of school quality**—teaching practice, classroom organisation & environment, and school leadership.

These gains will only be achieved by focusing the policy levers of government on the factors which matter most—**those which influence practice in the classroom**. However, these critical drivers are often the most removed from system level policy settings and levers. This means better evidence is needed—evidence which links practice at the school and classroom level to policy and performance. Against this backdrop, the role of government can be viewed in the following ways:

- **Effective system settings** are seen as a pre-condition to provide the environment for schools to identify and invest in effective practice in their own unique context.
 - These system settings, among other things, set standards and performance expectations for the teaching profession, establish the curriculum that educators use to guide their teaching practice, and guide the strategic focus of schools through processes of accountability for student outcomes.
 - At the educational front-line, quality improvements result from individual schools undertaking interventions to change their school practice and teaching practice to pursue school-level goals best suited to the unique characteristics of their student intake. These include measures like teacher professional learning and improvements in pedagogy, which are known to be the most important drivers of student outcomes.
- In this context, government can play a central role in **curating and evaluating the evidence base** which schools draw upon when making decisions about their practice and management.
 - Schools do not make decisions about professional learning and pedagogy in isolation. Indeed, they are influenced by a wide range of sources, such as professional bodies, private educational businesses, academics and government.
 - Current examples of best practice in collating such an evidence base include the NSW Centre for Education Statistics and Evaluation and the Education Endowment Foundation’s *Teaching and Learning Toolkits*.
 - These enabling initiatives can influence the consistency of practice and management interventions which schools utilise to improve their performance. A strong evidence base serves as a useful guide for policymakers and school decision-makers about interventions that improve student outcomes.

A high quality and transparent evidence base, when combined with **effective accountability processes**, may provide the necessary assurance that the schooling system is investing in higher quality education practice. One of the most important levers available to the Australian Government to change school practice is that of funding. One of the aims of the government’s *Quality Schools, Quality Reforms* initiative is to ensure “public accountability for the way in which funding is distributed, how that funding is used behind the school gate and achievement of outcomes”. This transparency – around the funding given to schools, the interventions being tried in schools, and the outcomes for students subject to these interventions – helps to continue the development of the education evidence base described above. This study has identified a significant and diverse range of schooling policy interventions which have occurred in Australian schools over the past 10-15 years. While diversity and complexity in policy design and application need not necessarily be a shortcoming, the lack of a consistent and universal basis for evaluating the impact of policy on student-level outcomes means there exists little capacity to ensure Australia is on a path towards overall school improvement.

It may be that Australia can make material progress in improving school quality not by making new or different interventions, but instead by more consistently adopting and applying proven best practice (and distributing resources accordingly).

This can only be achieved through better data and evidence, and better sharing of and access to that data, across Australia’s schooling system. Through greater transparency of student outcomes, school practice and system settings, government—in particular, the Australian Government—will be in a better position to evaluate current initiatives and practice across Australia.

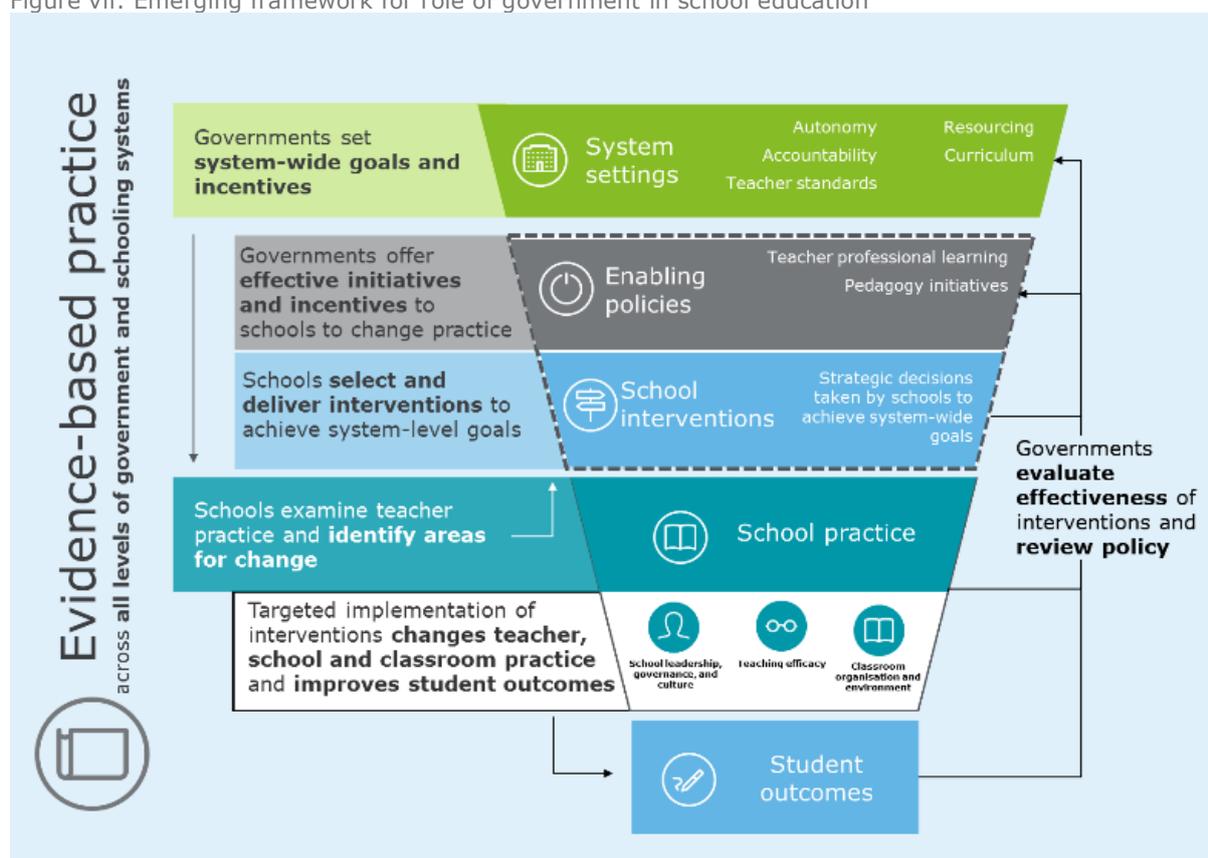
More specifically, governments can play a key role in the improvement of school quality through ensuring that policy makers consistently:

- Demonstrate the link to improvements in teaching practice before investing in initiatives intended to drive improvements in outcomes.
- Set a strategic and long-term focus on the outcomes impact of sustained changes to practice at the classroom level.

- Rely on, and encourage the use of, evidence-based interventions across Australian classrooms.
- Collect and share data and evidence on how interventions result in improved practice.

Figure vii below presents a framework for understanding the factors that influence school practice, and the role of government in determining system settings and enabling initiatives which provide the necessary conditions for schools to identify and invest in high quality practice. This framework emphasises the role of government in monitoring and evaluating the effectiveness of school practice to inform the system settings and enabling initiatives which guide decision-making, while simultaneously holding the system accountable for driving improvements in student outcomes.

Figure vii: Emerging framework for role of government in school education



Source: Deloitte Access Economics analysis of PISA and TIMSS data

Conclusions and future research directions

This study has demonstrated how the Department can use available evidence from PISA and TIMSS to identify the key drivers of student outcomes and school quality in Australia, and measure their relative importance. This empirical methodology, and the underpinning measurement framework, provides the Department with a detailed and impactful evidence base to inform future directions for government. In particular, it has answered the key research questions established for this study by providing:

- A framework and methodological approach which provides greater **clarity** in understanding the drivers of student outcomes and school quality in Australia;
- A robust, detailed and *Australian specific* evidence-base which builds **confidence** for policy-makers in understanding the key factors that contribute to student outcomes in Australia; and
- An emerging framework to support the **strategic focus** and direction for policy-makers when considering the role of government in driving improvement in Australia’s school education system.

The framework and evidence developed through this study may be used to inform frameworks for improving student performance to be used by Australian jurisdictions to demonstrate a link between new policies and the drivers of school quality. That is, the measurement framework and

evidence base developed here may assist policy-makers in providing evidence of a link between policies—including those relating to resourcing—and effective school and classroom practice, measured at the level of the classroom.

Future research may build on and refine this methodology towards providing further evidence and insight. This future research may expand on this study by, among other things:

- Expanding the scope of the empirical analysis to examine the effects of different drivers of school quality over time (for example, by mapping key PISA and TIMSS questions across years).
- Adding further causal structure to the empirical analysis to understand how different drivers of quality affect each other, and then subsequently drive student outcomes (for example, by estimating the link between school leadership and teaching practice).

Deloitte Access Economics

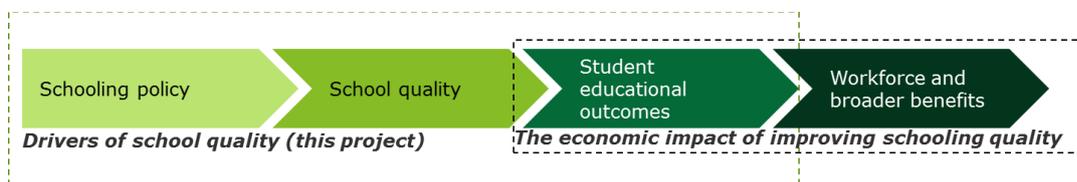
1 Introduction

Deloitte Access Economics conducted a research study in 2016 for the Department of Education and Training that filled a gap in the existing Australian literature regarding the quantitative impact of an increase in school quality (as opposed to school attainment) on the economy.

That previous study estimated the impact that an increase in school quality (associated with a sustained increase in PISA scores) had on total economic output. It demonstrated the mechanisms through which improved cognitive ability translates to higher educational attainment and accumulation of human capital, which in turn grows the economy through both higher investment and labour productivity.

This project builds on that research and seeks to complete the 'chain' which links the role of government, school leaders and teachers to economic outcomes, as illustrated in Figure 1.1 below.

Figure 1.1: Completing the link between schooling policy and economic outcomes



Source: Deloitte Access Economics

What is school quality?

In the context of this report, school quality is defined as **the contribution that a school makes to the outcomes of its students**, after controlling for contextual characteristics. The quality of a school is therefore a driver of student outcomes, as are other factors relating to the individual characteristics of students, or the context of the school and/or education system.

Drivers of school quality are defined as the **attributes of a school's practice and management that drive student outcomes**, and therefore determine a school's 'quality'. These attributes of practice and management are defined as factors over which a school has some degree of control—in contrast to a school's contextual characteristics.

Drivers of school quality are—by definition—also drivers of student outcomes. Indeed, the empirical models used in this study use student outcomes as a 'dependent variable' in each aspect of the analysis. Throughout this study, all empirical findings should be interpreted as drivers of student outcomes (that is, the effect of a given factor on students learning achievement, engagement or wellbeing). Findings which refer to the 'drivers of school quality' relate to those attributes of school practice and management which are found to drive student outcomes. It can be separated from other broad sets of factors that drive student outcomes:

- **Contextual factors**—including factors at both the student and school level (such as students' socio-economic status, school location, etc).
- **System level factors**—including characteristics of schooling systems, such as autonomy, accountability and resourcing, noting that these factors influence can influence school quality indirectly.

Studies such as Hattie's Visible Learning (2009) have sought to provide a comprehensive overview of each of the key factors that influence student outcomes in school education. However, a lack of evidence from Australia, and the piecemeal nature of research on particular aspects of schooling, conducted at different times and in different countries and schooling systems, make it difficult to gain a complete picture in an Australian context.

The PISA and TIMSS datasets provide an opportunity to bridge this gap by exploring the impact of multiple drivers of student outcomes within the same study. In addition to student assessment, PISA and TIMSS collect detailed information about the students and their school environment. Together these two datasets provide a rich source of information about school quality and differences between schools and between classrooms. Importantly, because these drivers can be tested within the same dataset (that is, with the same students and schools) they allow fully consistent and comparable effect sizes to be obtained.

The Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS)

PISA is an assessment of 15 year old students from different class years, classes and teachers. The school environment questionnaires have a greater focus on the principals' view of the school, including their view of school management, school level practices and their staff overall. This means that PISA is an ideal source of information on differences in school quality *between schools*, different student experiences across different years, and in different classrooms.

TIMSS is an assessment of selected year 4 and year 8 classes. The school environment questionnaires have a focus on what happens in a particular classroom from the perspective of students and their teacher. This means that TIMSS is an ideal source of information on differences in school quality from a *classroom* perspective.

The core contribution of this current work is through the way in which it matches the indicators contained in these datasets, to a measurement framework of the different drivers of school quality, which allows this empirical exercise to be undertaken in a systematic fashion.

The overarching research questions that this project seeks to address are:

- What evidence is available from the PISA and TIMSS datasets on the drivers of school and classroom quality and student outcomes in Australia?
- How might the content of these datasets be evaluated now and in the future?
- Using this data, how much do different aspects of practice in the school and classroom matter in determining both school quality and student outcomes?
- Based on this evidence, to what extent can the system improve student outcomes and school quality, using the levers of government policy?

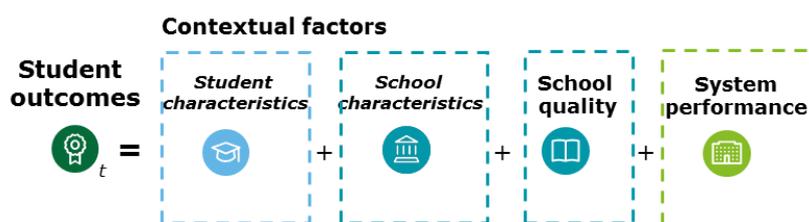
More details on the high level approach are contained below, and set out in full in subsequent sections of the report.

1.2 A measurement framework for drivers of student outcomes and school quality

Both the classroom drivers of outcomes, and the broader contextual drivers, have been extensively studied in the literature. This literature allows an identification of the key themes that drive outcomes at the different levels within the schooling system: namely, *contextual factors* beyond a school's control, *school quality* factors within a school's control, and *system level factors* amenable to policy at the jurisdictional, sector or national levels.

Each student's outcome can be taken as a combination of these drivers specific to that student. That is, a student's outcome depends on characteristics unique to them (such as their socio-economic status), the school they attend (such as characteristics of the student cohort), the quality of the school (such as characteristics of the teachers), and overall system factors, as illustrated in the stylised equation presented in Figure 1.2.

Figure 1.2: Thematic equation of measuring drivers of student outcomes



Source: Deloitte Access Economics

Measuring the drivers of outcomes at each of these levels is therefore important in fully understanding student outcomes. This lends itself to a regression based on a 'multi-level' model that explicitly accounts for the nested nature of different students within a school, and different schools within a schooling system. Such models are used to estimate the 'value-added' of schools, by identifying how much variation in students outcomes can be explained by the school a student attends, relative to the specific contextual factors that drives a student's outcomes.

A methodology for measuring the drivers of student outcomes, and school quality

To empirically estimate the relative importance of the different drivers of student outcomes it would be necessary to identify the extent of each driver faced by each student and compare that to the student's outcome (as measured, for example, by test scores). Naturally, it is not simple to identify and quantify the quality of the educational environment a particular student faces and this makes testing the size of the effects of these drivers on student outcomes difficult.

However, the PISA and TIMSS databases do contain questions that provide *indications* of these quality drivers. By carefully matching these indicators to identified drivers of school quality it is possible to undertake empirical analysis that identifies the different drivers of quality that have the greatest impact on student outcomes.

The key approach of this work therefore contains three research streams:

- A review of the literature is undertaken to identify the established drivers of school quality. This review leads to key themes that are known to drive quality, such as those that relate to teaching practice, school leadership, or school autonomy, for example.
- The indicators in the PISA and TIMSS datasets are then matched to these themes based on which theme each indicator is most likely to represent. A representative indicator (or set of indicators) is then selected in order to proxy and measure the extent of each quality driver (theme) faced by a student. This selection of representative indicators is based on a statistical selection process.
- Regression analysis is then undertaken to analyse the relative effects of the identified drivers on student outcomes, and to compare the relative sizes of these effects in order to determine what drivers of quality have the greatest effect on outcomes.

These observable characteristics at the classroom, school, system, and jurisdiction level are central to this study's enquiry into what drives student outcomes, from the perspective of government. In thematically constructing and identifying these drivers of student outcomes, the aim of this study is to—to the greatest extent possible—link these back to tangible levers for government, and to identify where governments should broadly direct their attention to have the greatest impact on student outcomes (and in what contexts and circumstances). This work does not go so far as to consider the impact of particular initiatives, but rather provides a broad strategic framework in which governments and schools have areas of focus in improving school outcomes.

The findings of this study are only accurate to the extent that the PISA and TIMSS questions are good proxies for the themes they are intended to represent. In many cases, judgement has been applied when interpreting and categorising these questions for the purposes of analysis and alternative approaches to categorisation may result in different empirical findings. This study looks to examine the effects of education practice on outcomes using proxy measures of outcomes

available in these datasets. This is in contrast to an approach built on actual field analysis of practice in classrooms, which may provide more nuanced and robust findings on the drivers of school quality in schools (albeit, for a smaller subset of observed schools and teachers).

In this regard, this study has sought to develop a robust and replicable methodology which may be used by the Department to explore the relative effects of different drivers of school quality using the PISA and TIMSS datasets. This would naturally be achieved by building on and refining this initial approach over time.

1.3 The structure of this report

The report is structured as follows:

- Chapter 2 provides a summary of the relevant literature, to identify the established drivers of student outcomes at the student, school, and system levels. Based on this summary, a measurement framework to underpin the empirical analysis is established.
- Chapter 3 summarises the approach taken to matching the findings from the literature review to the PISA and TIMSS datasets, and how this is then used to test empirically for the core drivers of students' outcomes.
- Chapter 4 presents the findings from the empirical analysis.
- Chapter 5 summarises the analysis and provides the key implications for government by the findings and identifies future research that could build on the results of this analysis.

2 The drivers of school quality – establishing a measurement framework

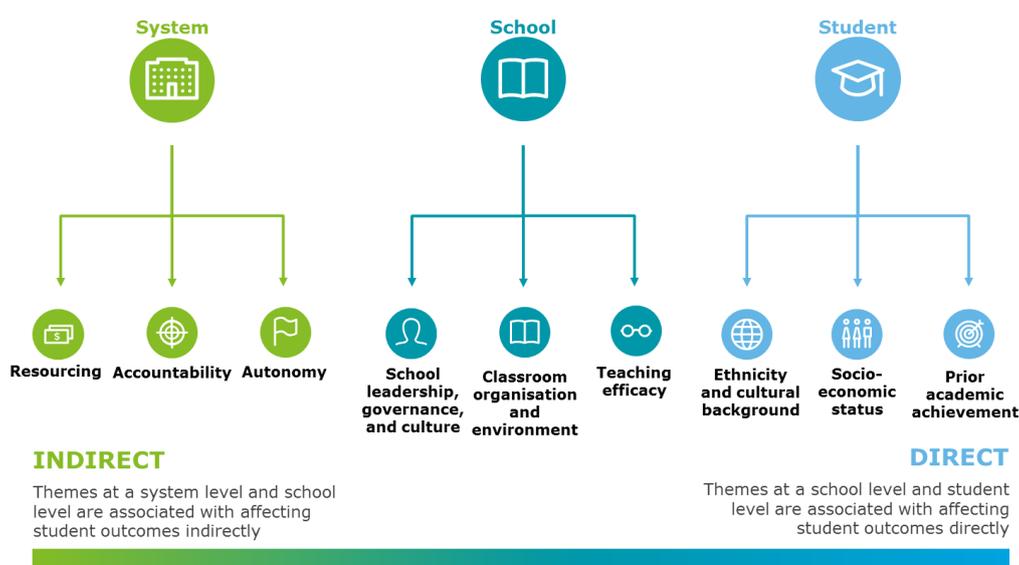
A plethora of literature highlights how different aspects of school quality influence student outcomes. *Visible Learning*, John Hattie’s synthesis of over 800 meta-analyses relating to student achievement, is perhaps the most notable collation of such literature.²

From a targeted review of this literature, a measurement framework was developed to identify, describe and evaluate key drivers of student outcomes, at the student, school and system level. Frameworks are a useful way of organising the literature into a logical, coherent and consistent basis for empirical measurement,³ where the different factors that drive student outcomes can be separately identified and measured. The framework is used to inform and guide the empirical analysis outlined in chapters 3 and 4 by providing:

- A common understanding of the terminology used in the study
- Identifying the relationship of each factor with student outcomes, including:
 - How each factor drives student outcomes;
 - Which drivers are contextual in nature and which are amenable to change by schools and by government; and
 - The links between different drivers of student outcomes.

The themes identified in the literature are organised below as *system-level*, *school-level* and *student-level* drivers of student outcomes. Within each level, the drivers of student outcomes are classified into nine anchor themes. A high level view of the measurement framework is introduced in Figure 2.1 below.

Figure 2.1: Overview of measurement framework



Source: Deloitte Access Economics

² Hattie (2008).

³ See, for example, Australian Bureau of Statistics (2015).

Each anchor theme has a number of sub-themes which were identified in the literature review as the most important areas to be considered when thinking about each theme. For example, within the school-level theme 'classroom organisation and environment', there are several sub-themes (as illustrated in Figure 2.2).

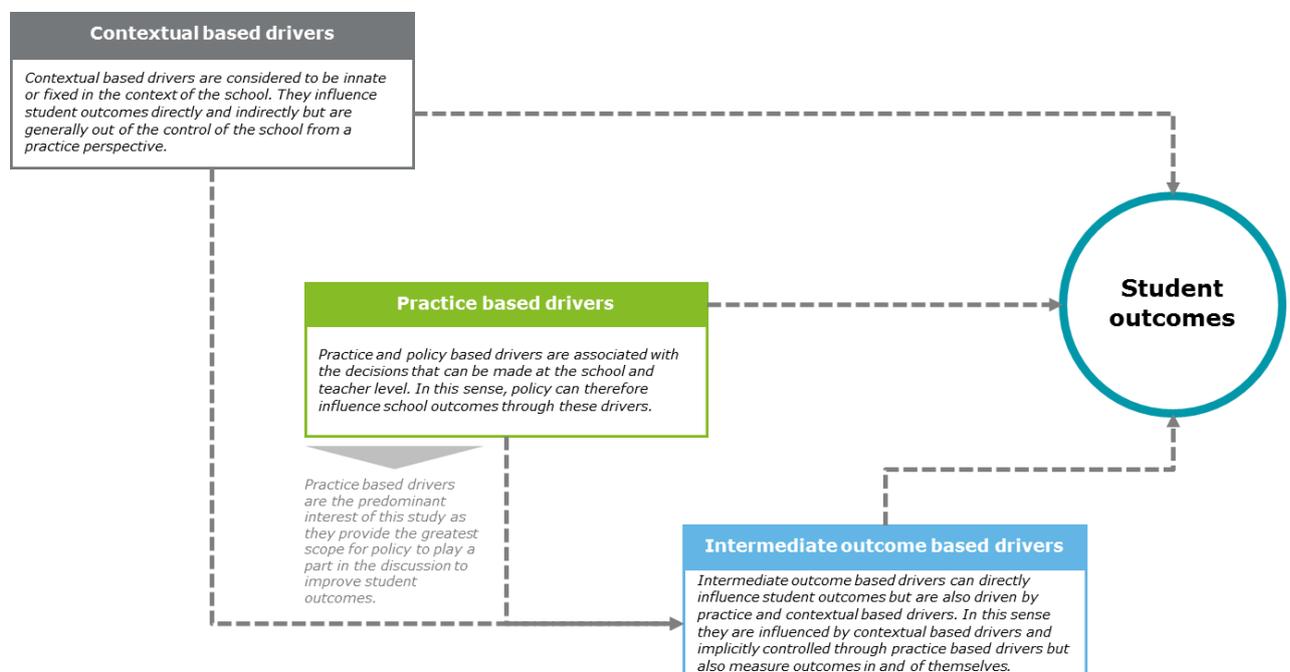
Figure 2.2: Example theme and sub-theme: classroom organisation and environment



Source: Deloitte Access Economics

It is important from a system and school perspective to understand how each of these themes and sub-themes influence student outcomes both directly and indirectly. As outlined in chapter 1, many determinants of student outcomes are *contextual* drivers – that is, outside of the control of the school. Some drivers directly determine student outcomes (*practice based drivers*) while other drivers are *intermediate outcomes* in and of themselves, that then contribute to students' academic achievement. Figure 2.3 below illustrates the conceptual link between these types of driver.

Figure 2.3: Overview of the conceptual link between intermediate outcome, contextual, and practice based drivers of student outcomes

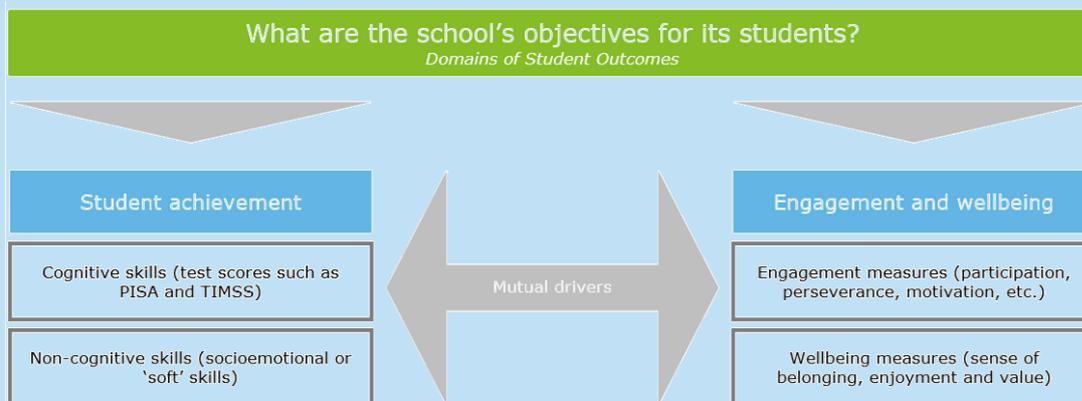


Source: Deloitte Access Economics

Ultimately, this report identifies the school- and classroom-level drivers of most importance and explore the role that government may ensure that ongoing system reforms are directed to the areas which make the most significant differences to outcomes.

Different domains of student outcomes

The notion of educational outcomes for students is varied. Broadly however, they can be considered to be captured by two domains, captured in the figure below.



Student achievement

The idea of student achievement as an outcome domain broadly captures the different skills associated with educational experiences of the students. A distinction is often made between cognitive skills, as outlined above, and non-cognitive skills, which can also be referred to as soft skills or socioemotional skills. These non-cognitive abilities relate to important personal attributes such as perseverance, motivation, self-control, conscientiousness, perseverance, sociability, and curiosity (Heckman, 2004; Heckman and Kautz, 2012). These are also considered elements of self-efficacy (considered further in section 2.1.1).

Non-cognitive skills are therefore significant and important educational outcomes that relate strongly to human capital. Indeed, Conti and Heckman (2014) conclude that conscientiousness is strongly correlated with attainment and labour market outcomes. These researchers argue that such traits are generally stable, but do evolve slowly over time and are malleable. As such, it has been found that formal education and other interventions can shape non-cognitive skills in such a way as to improve individuals' human capital over time.

This does not necessarily suggest that tests of student achievement are perfect measures of the elements of the cognitive and non-cognitive abilities that contribute to human capital, but they are nonetheless strong indicators, particularly in contrast to indicators of educational attainment. That is to say, attainment in and of itself does not capture all of the aspects of an individual's value to future employers, particularly relative to more nuanced measures of student achievement. While intuitive, this observation can have profound implications for public education policy centred on improving attainment in order to raise earnings (Burgess, 2016).

Student engagement and wellbeing

Beyond the measures of student achievement, the quality of a school can also be considered in terms of other forms of student outcomes, which include student engagement and wellbeing (Lamb et al., 2015).

Student engagement and wellbeing are important not only because of their relationship with student learning, but also because they represent a disposition towards schooling and life-long learning (Willms, 2003). Engagement and wellbeing therefore have more ontological implications for students beyond their schooling life.

In practice, the relationships between student engagement, wellbeing, and cognitive and non-cognitive achievement are highly complex and in many cases occur in a largely contemporaneous fashion, which limits the extent to which researchers can understand their causal inter-dependencies.

2.1 The factors which contribute to student outcomes – a brief literature review

To motivate the development of the measurement framework, this section presents a brief review of the international and Australian literature on the drivers of student outcomes and school quality.

2.1.1 Contributions from the student and home

Student level characteristics, including the student's home environment, are primary factors known to determine educational outcomes. These factors are generally not influenced directly by school practice—they are contextual factors which affect a student's learning, in the sense that a student brings them with them into the classroom.

The literature relating to contributions from the student and home identifies three core themes that influence student outcomes: self-efficacy; socio-economic status; and ethnicity and cultural background.

Student prior achievement and self-efficacy

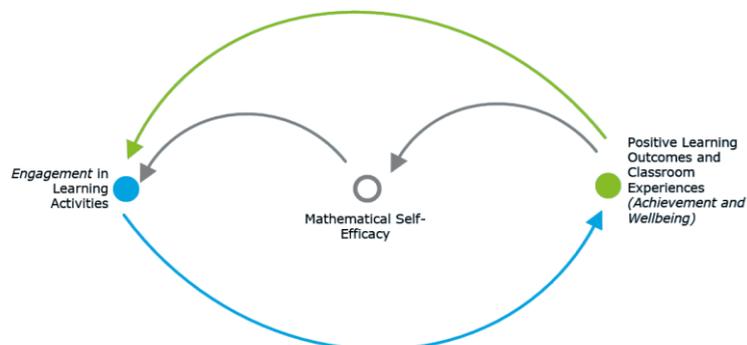
When considering the drivers of educational outcomes, *prior achievement and self-efficacy* play an important role in influencing student learning achievement. Student prior achievement and self-efficacy are also intrinsic measures of student *engagement and wellbeing*, which may be considered outcomes in and of themselves (Almlund et al 2011, and Cunha et al 2006).

Prior achievement, in part, reflects the existing abilities of students, A range of studies have demonstrated that prior school achievement (as a measure of a student's existing ability in a subject area) is a significant predictor of current achievement (see, for example, Hemmings, Grootenboer and Kay, 2010). However, this is inextricably tied to the overall attitude that students bring to learning in the subject area. Self-efficacy is a type of personal cognition defined as *people's judgements of their capabilities to organise and execute courses of action required to attain designated types of performance* (Bandura, 1997). Using the mathematics classroom as an example, Warwick (2008) establishes four main sources that influence students' judgements of capabilities and execution:

1. **Performance experience (prior achievement):** high performance in math assessments generally strengthens students' beliefs about their abilities in students whilst repeated low performance will weaken them;
2. **Vicarious experience:** the process by which a student compares oneself with peers;
3. **Verbal persuasion:** comments made by teachers (or parents) regarding the ability of a student to complete mathematical tasks;
4. **Psychological and affective states:** inner-feelings (*anxiety, worry, tension, confidence, happiness etc.*) that might be provoked by the student having to undertake mathematical tasks.

Self-efficacy is related to engagement in learning activities and subsequently is manifested through learning outcomes. Therefore, positive self-efficacy feeds into stronger levels of engagement and results in positive learning outcomes which feed back into engagement in learning activities both directly and indirectly (*through improved self-efficacy*).

The feedback loop is presented below:



Source: Deloitte Access Economics, adapted from literature discussed in this section.

Some researchers have suggested that teachers should pay attention to the development of students' perception of competence in addition to the level of actual competence itself, as Pajares (2003) notes:

"...there are situations in which inaccurate self-beliefs, rather than a weak knowledge base or inadequate skills, are responsible for students short-changing themselves academically."

At a foundational level, different aspects of self-efficacy are perceived as inputs which drive engagement and wellbeing and ultimately influence academic performance. For example, a student displaying a strong degree of confidence in learning is more likely to achieve better outcomes at school (Multon, Brown, & Lent, 1991). Motivation is an important related concept; a motivated student, with sufficient classroom autonomy, worthwhile goals and a supportive feedback system is more likely to achieve better educational outcomes (Dornyei, 2001).

The influence of prior achievement and self-efficacy implies that, all else being equal, investments in early education can be effective, through their effect on self-efficacy and by supporting student engagement and promoting wellbeing, and subsequently performance. In this sense, other intermediate outcomes, such as engagement and wellbeing,⁴ have risen to the attention of researchers, as they are considered objectives for schools in and of themselves.

When considering measures of school quality and performance, one must account for prior achievement as an explanatory factor when seeking to understand the value added to students' learning by teachers or schools (Lu and Rickard, 2014).⁵ In this regard, many measures of teacher and school performance focus on *learning gain*, not just absolute performance.⁶

Socio-economic status

From an empirical perspective, there is a strong link between student outcomes, and socio-economic status at a school and community level (Betts, Zau and Rice 2003). The impact of a student's socio-economic status on educational outcomes is significant and robust to jurisdictional contexts and across time.

At an aggregate level, Heyneman and Loxley (1983) find that socio-economic status is a more powerful determinant of achievement in higher-income countries, explaining 35% of the variance in student test scores. In an Australian context, the use of the family occupation and education indices in Australian jurisdictions like NSW and Victoria, or similar measures of parental education and/or occupational background, are positively associated with educational outcomes (Lu and Rickard, 2014).

⁴ Intermediate outcomes are outcomes which are driven by factors that affect school performance but are also drivers of school performance in and of themselves. For example, students are more likely to be engaged in school if they have teachers using effective practices; this engagement drives their performance in overall school outcomes measures.

⁵ It also must be acknowledged that school quality influences self-efficacy which in turn drives student performance.

⁶ In the measurement framework set out in this study, the impact of school and teacher quality is, in effect, assumed to be the 'learning gain' beyond a student's existing ability (or prior achievement), given a student's other characteristics. However, it is not possible to directly measure prior achievement in the context of PISA and TIMSS, as these are not longitudinal studies (they do not observe the same group of students in each testing period) and so these findings should be interpreted with caution.

Socio-economic levels and mechanisms affecting student outcomes

Socio-economic status may be more important at the school than at the individual level—particularly after accounting for the effects of prior academic achievement (which is strongly linked to socio-economic status). This suggests that the environment of the local community and the involvement of parents in education (particularly at home), while highly correlated with a student's socio-economic status, may be more causally significant than socio-economic status in and of itself (Hattie, 2013). In this sense it is not only the student's measure of socio-economic status that influences student performance, but the intermediate mechanisms (parental involvement, home learning environment, etc.) that also drive student outcomes.

The mechanics by which socio-economic status affects student outcomes is diverse and complex. For example, there is evidence to suggest that unobserved family practices may differ significantly by socio-economic status, including the nature of the home learning environment (including access to learning resources at home), the use of out-of-school time and parental involvement in learning (Belfield and Levin 2002). Differences at the home level is emblematic that it is the process by which socio-economic status affects student outcomes rather than the nominal socio-economic status of the student.

Organisations like the OECD (through PISA) and ACARA develop composite index measures which capture a diverse range of aspects of a students' background and home environment. These measures are typically calibrated to predict variations in student outcomes as accurately as possible, for the purposes of identifying relative educational need.⁷ In this regard, measures of socio-economic status in schools are practically developed as proxies for learning outcomes, based on observed contextual factors. This suggests a high level of endogeneity between measures of current achievement, prior achievement, and socio-economic background, which has important implications for measuring and understanding school performance and school quality in different contexts.

When comparing student outcomes on an international level, the reading outcomes of students from high socio-economic backgrounds do not substantially differ, as seen in the case of Canada and the US through PISA data (Willms, 2004). Students from low socio-economic backgrounds however, fare markedly better in Canada than the US, suggesting that system level policies may influence the level of equity in educational performance. Interestingly, the difference in the extent of socio-economic inequalities between Canada and the US stems mainly from differences *between schools*, indicating that school level interventions, especially at the low socio-economic end of the spectrum can be implemented to correct the inequity in schooling.

Ethnicity and cultural background

Contemporary schooling in Australia is generally of high quality. The dominant cultural expectations of the schooling systems, however, can negatively impact students from other ethnic and cultural backgrounds.

Indigenous students in Australia often experience poor student outcomes even after controlling for remoteness and socio-economic status (Biddle and Cameron, 2010). This discrepancy in student performance can potentially be attributed to the differences in learning approaches at a cultural level. As Nakata (2007) notes, knowledge transfer is deeply entwined in Indigenous cultural and social practice:

⁷ See the PISA index of economic, social and cultural status (ESCS).
<https://stats.oecd.org/glossary/detail.asp?ID=5401>

"the Indigenous epistemological bases of knowledge construction ... are embedded ... in ways of story-telling, of memory-making, in narrative, art and performance; in cultural and social practices, of relating to kin, of socialising children; in ways of thinking, of transmitting knowledge"

By not utilising these bases of knowledge construction in mainstream education systems, educational outcomes of students are impacted. In Australia, Indigenous disadvantage begins early in their schooling life and continues throughout childhood (Biddle and Cameron, 2010). There is a strong imperative to address the disparity in educational outcomes which arises from differences in culture or ethnicity.

The interactions between teachers and students of different ethnicities can result in differing educational outcomes, potentially as a result of bias in assessing student performance (Entwisle and Alexander, 1988). Student outcomes are therefore driven by not only their cultural and ethnic background directly, but also through the interactions with the school and teachers. Students who have migrated to Australia may have little to no prior schooling experience, which may have been in a different cultural setting to the Western classroom context. Students from culturally and linguistically diverse backgrounds may also need additional support to develop English language and learning skills (Australian Institute of Family Studies, 2015).

Governments, through their role in developing the education evidence base and setting standards for the schooling system, can therefore guide teachers and schools to establish approaches in accommodating students from different cultural and ethnic backgrounds. Resources such as ACARA's *English as an Additional Language or Dialect Teacher Resource* support teachers as they develop teaching and learning programs in the Australian Curriculum. One aspect of its development is to help teachers understand students' cultural and linguistic diversity and how this understanding can be used in the classroom.

2.1.2 Contributions from the classroom and the teacher

At the school level, several factors influence student outcomes. First is school leadership, governance, and culture, which can be considered a more indirect driver of student outcomes that effectively determines the conditions in which teacher and student interact. Second, classroom organisation and environmental factors are those that drive student outcomes directly through their influence on the classroom learning environment. Third, teaching efficacy, including pedagogy and curricula, encompasses the most direct and significant drivers of student outcomes.

School leadership, governance, and culture

School leadership is defined as the process of guiding and leading the talents and energies of teachers, students and parents toward achieving the educational aims set by the school. While there is no single model of the best practice of effective leadership, it is possible to identify a common set of broad educational values, competencies and strategic actions which result in improved leadership qualities in teachers and principals which lead to improved educational outcomes. In particular, school leaders can improve teaching and learning through setting a 'mission' for the school: by focusing on a small set of goals, directing staff attention to initiatives relevant to those goals, and taking a transformational approach to staff motivation, commitment and working conditions (Day et al., 2009). School leaders can improve outcomes not just by these influences, but by setting directions for a school's pedagogy, and classroom learning environment which can be enacted by teachers themselves (Marks and Printy, 2003). In other words, school leaders play a role in developing models of *distributed leadership* within schools, where teachers have flexibility adapt to the needs of the students in their classrooms.

Beyond school leadership developing the autonomy and capacity of teachers within the classroom, administrative support and guidance to principals and school leadership teams can be important in the short run. Evidence exists that in Australia, the ongoing work of administrating a school can inhibit school principals from focusing on the school's strategic direction (Watson, 2009). However long term leadership capacity building, to focus on a school's particular goals, is necessary to sustain improvement in educational outcomes. School leadership, governance and culture,

however, also covers the engagement of parents in the culture of the school. Significant evidence exists in the literature that “parents’ attitudes, behaviours, and actions in relation to their children’s education have a substantial impact on student learning and educational attainment” (ARACY, 2012). Engaging in their child’s school community permits parents to socialise their children for academic achievement, and to facilitate effective communication between schools and homes.

Ultimately, school leadership and governance exists to establish a positive school-wide culture that seeks to promote effective schools and improve student outcomes (Grant, 1988; Lightfoot, 1983). However, links between attributes of the school of this kind and student outcomes are generally uncommon in literature. Rowe (2003) cites substantive and methodological difficulties in connecting organisational factors with student outcomes. The substantive issue arises from determining the explicit link between school organisational factors and student outcomes. In light of this, other levers held by government such as school autonomy can play a part in enabling school leadership and governance to drive the initiatives required to improve student performance, as seen in case studies from Caldwell (2015). School leadership therefore broadly functions as an effective pre-condition that implicitly enables improved student outcomes to be achieved via the conditions in which teachers work and deliver classes.

Classroom organisation and environment

At the classroom level, drivers of student outcomes can broadly be grouped into three sub-themes.

Classroom context involves the structural or foundational elements of the classroom which are not determined by teachers, such as class sizes. There is little evidence in the literature to support the notion that reducing class sizes is an effective method in increasing student achievement. Ehrenberg and Brewer (2001), Hanushek (1999) and Mishel & Rothstein (2002) all find that although hundreds of studies have looked at the impact of reducing class sizes, the vast majority have found small or inconclusive benefits, despite the significant costs associated with increasing the ratio of teachers to students in education systems. Remoteness of a school has also been observed to have a negative impact on outcomes in both Australian and international contexts (see, for example, Gonski, 2011). Beyond the increased cost pressures associated with regionality, the literature has suggested that the “...homogeneous character of ... neighbourhoods creates enclaves denying young people social and cognitive challenges” (Sellström, 2006).

Aspects of classroom organisation – namely, strategies pursued by the school to shape the way students and teachers interact in the classroom – can have beneficial effects on student outcomes. Effective teachers use rules, procedures and routines to direct students’ attention towards their learning, and to avoid distractions (Stronge, Tucker and Hindman, 2004).

Classroom environment refers to the resultant climate of the classroom as a function of the previous two sub-themes. Peer culture, or how students interact with each other, influences the development of the classroom environment while also being a product of the context of the classroom and how it is organised by the school. The effect of peer culture is difficult to credibly estimate, given the possibility that individuals are likely to self-select into peer groups based on unobservable characteristics. Despite this, Hoxby (2000) finds that when there is a change of 1 percentage point in the reading scores of an individual’s peers, that individual’s reading scores improve by between 0.15 and 0.4 percentage points, suggesting there are some peer ‘spillovers’ toward improved educational performance. Evidence also suggests that these impacts are stronger within ethnic groups, and that these effects are likely to diminish as peer groups become more similar. In this sense, the interaction between ethnicity and peer effects influences educational performance depending on the ethnic composition and degree of ethnic diversity within peer groups.

Ability grouping provides another example of a common technique by teachers that ultimately affects the classroom environment in which they teach. Burke and Sass (2013), with controls for differences in student ability and teacher effectiveness, find evidence for peer effects at the classroom level, but not at the grade level—suggesting that the organisations of students within classes can help to redress the negative effects of inequitable peer groupings. Importantly, grouping students into different classes based on ability has minimal effects on student outcomes

and profound negative equity effects for students in 'low track' classes, though it may have positive outcomes for gifted students (Hattie, 2013). In contrast, the literature is inclined to support approaches that balance ability grouping *within* classes, as opposed to *across* classes, as it creates positive spill over effects, especially for low ability students.

Teaching efficacy

Hanushek, (2011), in a study on the impact of teaching quality on student outcomes, concludes that:

"No other attribute of schools comes close to having as much influence on student achievement"

The nature by which teaching quality influences student outcomes is complex in that there is no one-to-one causal relationship; rather, there is a complex interplay between different conditions and methods that produce varying student outcomes. When considering the effect on student outcomes, there is a large disparity between the most talented teachers and the teachers at the bottom end of the distribution. Due to this, and the large costs associated with other approaches (such as reducing class sizes), teaching quality has become the focus for much of the research in primary and secondary school education around the world (Hanushek, 2011).

Teaching quality can be conceptualised in several ways. Measures of teaching quality can include attributes that are both observed and unobserved within the classroom. That is, while it is clear that different teachers can lead to different educational outcomes, it is not always clear what the exact attributes of teachers are that lead to this. Indicators observable to administrators may include the teacher's level of academic proficiency (measured using observations of professional qualifications). Indicators that may be more difficult to directly observe may include teaching styles and other personal attributes that may be confounded with pedagogical practices that are not necessarily unique to the individual teacher.

In practice, both observed and unobserved attributes of the teacher matter, but when considering the implications of increased teaching quality it is important to understand the distinction between inherent attributes of the teacher (like latent ability and cultural orientation) and practice based attributes (like teacher development and instructional approaches) (Deloitte Access Economics, 2014).

- **Teacher attributes** – Static and indirect in the way it influences teaching efficacy, these drivers are present with teachers before they enter the classroom setting.
- **Teaching practice** – Dynamic in the context of the classroom environment, these drivers are direct in that they are transmitted immediately from the teacher to the student.

This study uses the term 'teaching efficacy', broadly synonymous with teaching quality, to characterise these attributes and practices which influence student outcomes.

Teacher attributes

Teacher qualifications can influence the quality of teaching practice through the skills and knowledge that teachers bring into the classroom. Betts et al (2003) find that teacher qualifications matter more in upper grades than in lower grades. In particular, teachers with a Masters qualification are estimated to increase the rate of learning among high school students by 20% and teachers with a Ph.D. are estimated to increase the rate of learning by 80%, suggesting that the level of qualification (as a proxy for teaching efficacy) matters for the rate of learning for the student. This suggests that having higher qualifications allows teachers to bring either domain specific knowledge or pedagogical expertise from their own learning experiences to influence educational outcomes.

Beyond degree qualification, strong attributes associated with content knowledge and subject matter expertise are indicative of positive teaching efficacy. The most effective teachers have a deep understanding of the subjects they teach, and if this deep knowledge falls below a threshold level, it can have detrimental effects to students' learning (Coe, Aloisi, Higgins, and Elliot Major, 2014).

Teaching practice

Teacher wellbeing and engagement is an essential pre-condition that provides strong self-efficacy for teachers to execute their classes effectively and sustainably.⁸ For example, teacher beliefs can include the teacher's reason behind adopting particular practices, the purpose they aim to achieve, the teacher's idea about what learning is and how it happens, and their conceptual model of the nature of teaching in the learning process (Coe, Aloisi, Higgins, and Elliot Major, 2014). Other forms of wellbeing such as teacher professionalism builds positive teaching efficacy in the classroom, this includes participation in professional development, supporting colleagues, punctuality, preparation, and liaising and communicating with parents.

Underpinning the other core element of teaching efficacy is the practice that goes on in the classroom. Through their practice, teachers have the ability to create a conducive learning environment to improve student outcomes by focusing on the following elements:

- Curriculum;
- Lesson planning and collaboration;
- Instructional approaches;
- Targeted teaching strategies;
- Student assessment; and
- Student feedback, assistance and development.

These are broadly the 'levers' teachers control during the teaching process that drive student outcomes. These approaches to teaching and learning are however, predicated on the notion that teachers possess good classroom management skills, which includes efficient use of lesson time, coordinating classroom space and resources, and managing behaviour with clear rules that are consistently enforced. These are the environmental components necessary for good learning rather than the components that students experience directly during the class (Coe, Aloisi, Higgins, and Elliot Major, 2014).

For example, Curriculum programs vary in effect and nature. At a high level, there are curriculum programs aimed at catering for gifted and disadvantaged students. In this regard, Hattie (2009) analyses three different curriculum effects:

- Accelerated: acceleration through the curriculum;
- Enrichment: enrichment of educational lives of some students; and
- Ability: ability grouping within the same curriculum framework).

The most effective method is to accelerate students through the curriculum, with enrichment and ability grouping being secondly effective and least effective respectively. It is important to note that the practice of curriculum programs by teachers in the instances above are often coupled with classroom organisation strategies (such as individually accelerating students and ability grouping). With regard to disadvantaged students, developmental curriculum programs like 'repeated reading', 'vocabulary programs', 'creativity programs' and 'phonics instruction' are found to have the greatest impact on student outcomes, with improvements being manifested through effect sizes of around 0.65 standard deviations of improvement, which ranks among the most significant in terms of student outcome effects (Hattie, 2013). From a school perspective, allowing teachers to tailor the curriculum for students of different backgrounds allows students to develop based on their educational needs (Goss et al., 2015).

⁸ Self-efficacy in teachers aligns closely with the same mechanisms that allows student self-efficacy to translate into positive student outcomes

The Sutton Trust Educational Endowment Fund

In the UK, The Sutton Trust Education Endowment Fund (EEF) Teaching and Learning Toolkit (TLT) provides a summary of educational research provided for guidance for teachers and schools on how to improve the attainment of disadvantaged pupils. Programs focused on adjusting curricula to be better suited to the learning style of students were found to improve student educational outcomes by around 2 months at very low cost.

The EEF has undertaken a number of studies of particular curricula programs, evaluating their effectiveness as drivers of educational outcomes and the relative cost of their implementation. One example is the Catch Up® Numeracy program, a one to one intervention for learners who are struggling with numeracy. The approach is based on research indicating that numeracy is not a single skill, but a composite of several relatively discrete component skills. The program was found to have a strong and significant effect size on numeracy outcomes, equivalent to around 3-4 months progress in student achievement. This was achieved at an estimated cost of £130 per pupil.

Large gains in educational outcomes can be made via direct pedagogical teaching strategies (Coe, Aloisi, Higgins, and Elliot Major, 2014). The quality of *instructional approaches* of a teacher which includes techniques such as effective questioning, reviewing previous learning, providing model responses for students, giving adequate time for practice to embed skills securely, progressively introducing new learning material (known as scaffolding), reciprocal teaching, and direct instruction are shown to be the more effective methods in improving student outcomes (Hattie, 2013). Professional development of this type have been highlighted as a major imperative for the Australian educational system (Jensen, 2010).

Lesson planning and collaboration strategies such as sharing by teachers of their conceptions about what constitutes progress through the curricula is essential to achieving improved student outcomes (Hattie, 2013; Gonski et al., 2011). These tools and methods can then be taken into the classroom to provide teachers with effective classroom practice on an individual level.

Overall, this evidence suggests that variations in teacher quality outweigh the impact of other education initiatives aimed at improving student outcomes (Jensen 2014; Burgess, 2016). Studies by Jensen (2010) have demonstrated that as much as a 10% increase in overall teacher effectiveness can be associated with a 19 point increase in PISA scores. In driving improvements to student outcomes across Australian schools, classroom quality is important not only on its own, but also to the extent that it interacts with a suite of other factors influencing educational outcomes (Jensen, 2010).

Teacher development

Learning and development and maintaining professional standards are flagged as essential elements in improving teaching quality (CESE, 2017). Teaching efficacy can be improved through approaches focussing on the wide range of in-class approaches. Generally however, initiatives geared toward improving teaching efficacy are cumulative in nature, in that students who have consistently effective teachers are more likely to get ahead (Sanders and Rivers, 1996). Strategically focusing on early intervention, and focusing on practices backed by a strong evidence base is paramount in effectively driving student outcomes.

Summary of the school level themes

The literature indicates that *school leadership, governance and culture* at a school level influences student outcomes insofar as they effect the classroom learning environment and the quality of teaching. While *teaching efficacy* is understood to be the most important (non-contextual) driver of student outcomes, the factors which influence teaching efficacy are complex and multi-faceted. Broadly, the literature indicates that teaching efficacy can be conceptualised broadly by two themes—teacher attributes and teaching practice.

Within the notion of teacher attributes, are characteristics of the teacher that are not directly associated with the classroom, including teaching specialisation and experience levels. Conversely, teaching practice concerns itself with the professional attributes relating to (i) teacher wellbeing and development and (ii) approaches to teaching and learning—these help to characterise what high quality teachers should be focussed on attaining. The component factors of school quality indicate the themes which this study will consider in identifying the key drivers of school quality within PISA and TIMSS.

2.1.3 System level factors

Although not at the forefront of the classroom experience, system level factors drive student outcomes by creating favourable conditions for schools to develop improved educational outcomes. Improvements in student outcomes given current resource levels, requires a mix of system and school level changes which are predicated on some degree of autonomy and accountability. A broad set of performance improving interventions that views education as a system, not as disconnected schools and sectors, is a critical component of this (Bracks, 2015). The major components that can be considered as important system level factors are: resourcing, accountability, and autonomy.

Resourcing

Literature and evidence on the relationship between school resources and educational outcomes consistently concludes that, *resourcing (and implicitly, 'funding') can be necessary but is not sufficient in its own right to improve educational outcomes* – that is, how resources are used is the critical issue, provided that, as a threshold requirement, the level of resources are adequate.

To elaborate, Odden et al. (2007) argue that schools require both *adequate* resources combined with an *effective use of funding* in order to improve performance. Darling-Hammond (2010) explores this theme, suggesting that how resources are used is critical to their impact on outcomes, as not all investments of resources have equivalent results:

"An effective system should create both a means for determining and funding adequacy and incentives to increase the likelihood of funds being wisely spent. At a minimum, states should not force schools to waste scarce resources through ill-conceived requirements."
(Darling-Hammond 2010)

Indeed, Hanushek, who has consistently argued that there is no systematic relationship between funding and outcomes, has acknowledged that:

"...there clearly are situations where small classes or added resources have an impact"
(Hanushek 2003)

As a result, when considering the nature and form of methodologies or models that allocate resources to schools within a system, the manner in which this methodology requires or enables resources to be deployed towards 'effective investments' is critical to the overall efficacy of the resourcing arrangements, and education system more broadly. Indeed, the findings from recent reviews of school funding arrangements in Australia, including Gonski (2011) and Bracks (2015), suggest that the level of funding and effective use of additional funding is critical to improving both the overall level and equity in students' educational outcomes.

Evidence on how resourcing drives educational outcomes can be broadly decomposed into two themes – Material based resources (the physical resources available to the school, such as books) and curriculum and practice based resources (the overall capacity of the school and toolkit of techniques available to be used in teaching particular students).

Material based resources

Analysis of Programme for International Student Assessment (PISA) data has shown that there is a correlation between the level of material resources (defined as the average adequacy of a school's educational resources) and mathematics performance, however the inequity in the allocation of material resources (defined as the difference in the quality of schools' educational resources between socio-economically advantaged and disadvantaged schools) is negatively correlated with mathematics performance—that is, both the adequacy of funding and the way in which it is allocated is important in driving student outcomes (OECD, 2013).

Generally, literature has indicated that funding has the highest impact when it is directed towards changes to teacher practice, but these effects vary on the basis of individual school and student context. System level contextual factors such as autonomy and accountability also contribute to the efficacy of funding and resourcing. However, resourcing is not as simple as '*more funding means improved student outcomes*'. Rather, it is more important to consider the manner in which funding is used.

Curriculum and practice based resources

There are a vast number of different programs available to schools and teachers to support the transfer of particular skills to students. An extensive amount of literature exists on the effectiveness of these different programs. As one example, there has been considerable debate around the use of phonics programs versus whole language courses of instruction for basic literacy skills (see, for example, Rowe, 2005).

In some cases, these programs might be directed at particular groups of students: for instance, literacy programs may be tailored for students learning English as an additional language or dialect. However, these varied school level programs can have very different impacts on student outcomes, and can have significantly variable levels of associated cost. An effective schooling system should aim to develop the institutional knowledge of schools such that they have access to evidence-based curriculum and practice based resources, and that they can select resources that help them to target their teaching approaches towards particular students.

Recognising the importance of understanding the efficacy of different programmatic investments in school quality, Evidence for Learning's *Teaching and Learning Toolkit*⁹ presents a summary of educational research to guide teachers and schools on how to improve the attainment of disadvantaged students.

⁹ Modelled after the UK Sutton Trust Education Endowment Foundation (EEF) *Teaching and Learning Toolkit*.

Examples from Evidence for Learning’s Teaching and Learning Toolkit

The research in the *Teaching and Learning Toolkit* is intended to be used to ensure maximum return is achieved for government spending on education resources. In particular, it focuses on how schools can make informed decisions with regards to spending on individual students, especially in attempts to raise achievement among disadvantaged students¹. Importantly this research considers both the relative impact and cost of different spending initiatives, indicating areas where increased funding can be linked with improved outcomes via particular drivers.

The results from the research presented in the tool kit show that, while interventions like reducing class sizes and improving parental involvement have similar impacts on student outcomes (increasing student attainment by around 3 months), they do so at very different costs, with class size reduction having a very high cost compared to a moderate cost for parental involvement.

Programs focused on pedagogical practices and curricula programs, like meta-cognition and self-regulation, and feedback mechanisms, have the highest impact on educational outcomes, for relatively low cost. This demonstrates the importance of programs like the Smarter Schools National Partnerships (SSNP) in providing funding for classroom drivers that have a demonstrable impact on educational outcomes and that can be implemented in a cost effective way (Parkville Advisory Group, 2014).

Selected Approaches	Average cost	Evidence security	Months impact	Summary
Feedback	\$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 8 Months	High impact for moderate cost, based on limited evidence.
Meta-cognition and self-regulation	\$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 8 Months	High impact for very low cost, based on extensive evidence.
Parental Involvement	\$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 3 Months	Moderate impact for moderate cost, based on moderate evidence.
Reducing class size	\$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 3 Months	Moderate impact for high cost, based on moderate evidence.

Source: Evidence for Learning Toolkit - See Appendix for full table

Even when drivers are demonstrated to have a significant positive effect on student outcomes, as found in the literature, the individual context of the school and classroom will still have a significant impact (EEF, 2014). This may mean that inferences drawn from meta-analyses and other syntheses of the literature are not applicable in certain contexts, highlighting that these results are to be reflected on in general contexts.

Therefore, contextual factors like autonomy, accountability and school leadership can be important in ensuring that funding is tailored to meet the needs of students in different classroom contexts, by supporting and enabling the drivers of student outcomes that have been evaluated as being most effective at a classroom level.

Accountability

Accountability is the broad notion that schools, districts, educators, and students are held responsible for their educational outcomes.

Generally, students in countries with externally administered exit exams consistently and substantially achieved higher scores as measured by internationally standardised exams (Hanushek and Woessmann, 2014). Externally administered exams can have strong impacts on the level of achievement, with the impact on average being more than a year of schooling. These results are robust to within country studies (where some regions have external exams while others do not, or some subjects are tested externally while others are not), suggesting that this is not driven by cultural differences. This reinforces the role of government in influencing student outcomes by administering external examinations for students.

At a school level, student achievement can increase as principals and teachers are held accountable because they are continuously subject to monitoring and review processes.

Accountability - school and system performance

Schools within Australian educational jurisdictions on average are afforded a relatively high level of autonomy, however this occurs within a robust framework of accountability. In Victoria for example, at least once every four years schools are subjected to a review within a system framework (DEECD 2013:10).

At a school level, review processes include (Caldwell, 2015):

- **Peer Review** - school management and externally accredited reviewers assess school performance leading to develop new four-year school plans
- **Exemplary Practice Review** – identification of exemplary practice in specific fields of practice are documented so they may be shared with the broader learning community
- **Priority Review** – an accredited review team undertake in-depth diagnosis of the causes underlying the school's low performance, with feedback being provided back to the community
- **Support and intervention Review** – Given the findings and feedback from the *Priority Review*, a design team will support to develop an intervention program which is then monitored by external bodies (e.g. Regional Services Group in Victoria)

Peer and Exemplary Practice reviews generally occur in high performing schools, whilst Priority and Support and Intervention reviews occur in low performing schools. In this sense, accountability frameworks not only hold schools responsible for their decisions in practice, but also serve as a means by which new and effective practices can be derived in light of performances of exemplary schools. Mechanisms for accountability can therefore promote varying conditions and structures by which schools can align and develop positive educational outcomes.

Appropriate accountability, both at a system and school level can play a pivotal role in directly improving student outcomes, and also indirectly providing a conducive environment for schools, educators, and students to drive improvements in student, engagement and wellbeing.

Autonomy

Autonomy in the schooling system is when schools have been devolved discretion towards the practice and management of schools. In Australia, schools across states and territories have been moving towards more independent and autonomous educational models that seek to improve student results through a more flexible and needs-based system.

Through analysis of PISA scores both internationally and in an Australian context, higher levels of autonomy are associated with higher levels of student achievement, provided there is a healthy balance of accountability (Hanushek and Woessmann, 2014; Caldwell, 2015). Higher levels of autonomy should align with strategies that are linked (via research and evidence) with professional practice and subsequently to gains in student achievement directly and not on structural changes for their own sake. Much of the evidence in aligning autonomy and professional practice strategies shows that school leadership and management can improve student outcomes: through building professional capacity (staff selection, professional development and appraisal); through the communication of purpose, process and performance at a school level, and setting priorities about performance based on data and evidence. Professional standards at a school management level and coherence and understanding between school management and educational practices compound the positive effects of autonomy.

Indeed, the different educational settings that exist within and across jurisdictions have influenced the degree of autonomy provided to schools over time. Over the past four decades, Australia has experienced a higher degree of autonomy within schools across all jurisdictions, further

highlighting the attention policymakers have put on differences in school environments, and not necessarily on the jurisdictional differences themselves (Caldwell, 2015).

Case study: Broadmeadows Primary School

Broadmeadows Primary School, in the northern suburbs of Melbourne, has had a long standing interest in higher levels of school autonomy. With the introduction of the Schools of the Future program in the mid-1990s, the school was one of the first to be selected providing a higher level of school autonomy.

Strategies such as internationalising professional development, training staff to mentor younger teachers, and allocating \$300,000 to employ a full-time curriculum coach are some of the autonomous strategies being implemented to improve teaching practice. Autonomy in building staff capacity has enabled positive school outcomes: in the 2014 NAPLAN tests, Broadmeadows scored higher on average in Year 3 and 5 literacy and numeracy compared to similar schools.

Autonomy has also enabled the involvement in other initiatives, such as the Powerhouse Schools project (Social Ventures Australia). This included initiatives such as: teachers being trained in the use of research-based learning behaviours toolkit, 30 percent of students being interviewed, with individualised programs being subsequently implemented, and parents attending forums to build awareness and capacity to improve their children's learning behaviour (Caldwell, 2015).

This robust body of evidence provides an example of how affording schools greater autonomy channelled through school leadership and professional practice can positively influence student outcomes provided there is sufficient accountability within the system. Governments can therefore set positive foundations for schools to lead and navigate their own professional practice needs to achieve better educational outcomes.

Summary

While system level factors can play a key role in driving student outcomes and school performance, the literature indicates that this is achieved through indirect transmission mechanisms, which support the pre-conditions and incentives for schools to invest in good practice and management, which drives high quality teaching and improved school performance. Overall, the literature indicates that how resources are used is much more important in determining student outcomes than how much funding is provided to schools. In this context, school autonomy, when supported by effective school accountability, is a powerful enabling force for effective practice and improved performance.

Critically, autonomous schools must have access to robust and transparent evidence on effective practice, and the capacity to be able to use data and evidence to target teaching strategies towards the needs of students in their individual context. These themes have strong parallels with aspects of reform which Australian schools have experienced over recent years, which are further uncovered and analysed as part of this study.

2.1.4 Implications for this study's methodological approach

This literature review has provided a thematic summary of the range of factors known to influence student outcomes, and a measurement framework underpinned by literature which serves as a basis to extract meaningful information from the PISA and TIMSS data sets.

The overarching theme that this literature review has highlighted is the importance of teaching efficacy, and, by association, school quality. Interconnected with this notion are more indirect drivers such as autonomy and strong school leadership and governance, which serve as pre-conditions that have lead-on effects for overall teaching efficacy. The purpose of this study is to further uncover the specificities of school quality present in the Australian education system using

the PISA and TIMSS datasets, and measure the extent to which variation in practice influence student outcomes. The ability to see which elements of school and classroom quality influence student outcomes can guide a more robust discussion in government about its role in the schooling system.

In parallel, it is also important to understand how other factors also interact with each other and how they contribute to student outcomes. The reviewed as part of this study highlights different scenarios to which certain factors influence student outcomes, which emphasises the importance of classification and measurement further outlined in Chapter 3. The empirical analysis will therefore include a wide range of factors outside of the teacher's control to (i) isolate the notion of school quality; and (ii) provide potential insights into factors that are out of the direct control of government.

2.2 How school performance and school quality is measured

As background and to inform the model developed for this study, outlined in Chapter 3 of this report, this section briefly discusses different approaches to measuring school performance and school quality in the literature.

Early studies in measuring school quality and educational outcomes (Coleman, et al., 1966; Jencks et al., 1972) suggest that school effects have little impact on students' learning outcomes, rather ethnic and SES background factors constituted the predominant effects of students' educational outcomes. The consensus around these findings at the time were consistent with social and political opinion (Rowe, 2003).

Rowe (2003) further highlights that the earlier findings have undergone much scrutiny on the primary basis of not accounting for the inherent nested nature of the educational system. This criticism spurred a range of studies into school effectiveness. Studies by Brookover, Edmonds, and Rutter began to analyse the contextual features of schools where students were performing better than their counterparts in comparable schools, after adjusting for the effects of intake characteristics. These early works found effective schools are characterised by a 'culture' oriented towards learning (e.g. professionalism amongst teachers, high expectations), and educational leadership (e.g. principal establishing agreed goals, increasing staff competence and involvement).

A commonality of these early studies is that they also lacked methodological sophistication required to model and analyse the complex interrelationships between inputs, processes and outcomes, including indirect effects and reciprocal effects (Rowe, 2003). Absent also, was the ability to take into account the inherent nested nature of the schooling system – that is schools organised students into classes taught by teachers (Rowe, 2003). Raudenbush and Wilms (1991) note:

"An irony in the history of quantitative studies of schooling has been the failure of researchers' analytic models to reflect adequately the social organization of life in the classrooms and schools. The experiences that children share within school settings and the effects of these experiences on their development might be seen as the basic material of educational research; yet until recently, few studies have explicitly taken account of the effects of particular classrooms and schools in which students and teachers share membership."

Rowe (2003) outlines two methodological advances worth noting. The first of which is the development of structural equation modelling techniques that enable the simultaneous estimation of interdependent effects among variables within a framework that takes into account measurement error, and the structural prediction residual. The second of the developments is the multilevel analysis methods that control for the inherent nested structure of the data, which enables the measurement of variables' effects at different levels of analysis (student, school, etc.).

These statistical developments have been implemented in value added models, both in Australia and internationally. The term 'value added' does not have a uniform definition, however, it is most commonly used to describe the additional value schools bring to the learning outcomes of students, after controlling for students' characteristics and attributes (such as prior educational

achievement) (DEECD, 2007). Hill (1995) distils the notion of value added as an attempt to "indicate the educational value that the school adds over and above that which could be predicted given the backgrounds and prior attainments of the students within the schools". The term 'school effects' or 'school performance' within the literature is equally interchangeable with the term 'value added' (Lu and Rickard, 2014).

Internationally and within Australia, value-added models vary in the technical approach to measurement as well as in the compositions of the model. The range of approaches reflect the different times and purpose when they were developed. Lu and Rickard (2014) outline four potential factors toward the difference in modelling choices:

- The policy and political environment could dictate the selection of certain variables and controls;
- Statistical models available at the time;
- Data availability; and
- Different views on the definition of school effectiveness and the underlying assumptions about students' learning growth.

There are three predominant statistical approaches for value added modelling:

- **Gain core model** – relatively simplistic in nature, this approach uses the average of the gain scores (the difference in pre-test scores to post-test scores) across all students in a school. Other variations include comparing the growth percentile for each students by taking the difference of the growth of a student and average growth of all students in the district/system with similar scores in previous tests. The growth measure is then aggregated across all the students in the school as indicator of the value a school adds to students' learning. A key assumption is that the effect of external and contextual factors (e.g. student SES) on students' achievement has been accounted for in their prior achievement. This assumption is not always met indicating limitations in the simplicity of the approach.
- **Covariate adjustment model** – a single level regression model, either at the student or school level. Using the current test score as the outcome variable, the covariate adjustment approach attempts to explain the variation in this variable by using a simple statistical regression method to account for prior achievement of students attending a school and other student and/or school level contextual factors. When controlling for such factors, the difference between actual school mean performance and predicted mean performance is then associated as an indicator of school effectiveness. One of the main drawbacks of this model is that it does not take into account the nested nature of educational data.
- **Multilevel models** –the multi-level modelling approach has strong advantages as it does account for the nested nature of educational data (Lu and Rickard, 2014). The degree of levels that can be simultaneously modelled are:
 - Intra-student level: performance across tests in current and previous years;
 - Student level;
 - Teacher level; and
 - School level.

Previous multilevel models have attempted to approach measuring school effectiveness by omitting non educational factors, with the assumption being that prior achievement has accounted for this as proposed by Sanders (2000). Other studies have found insufficient evidence supporting this claim (Griffin, Woods & Nguyen, 20005; OECD, 2008), leading to the inclusion of non-educational factors. The resultant model has been labelled as *multilevel contextual value added models*.

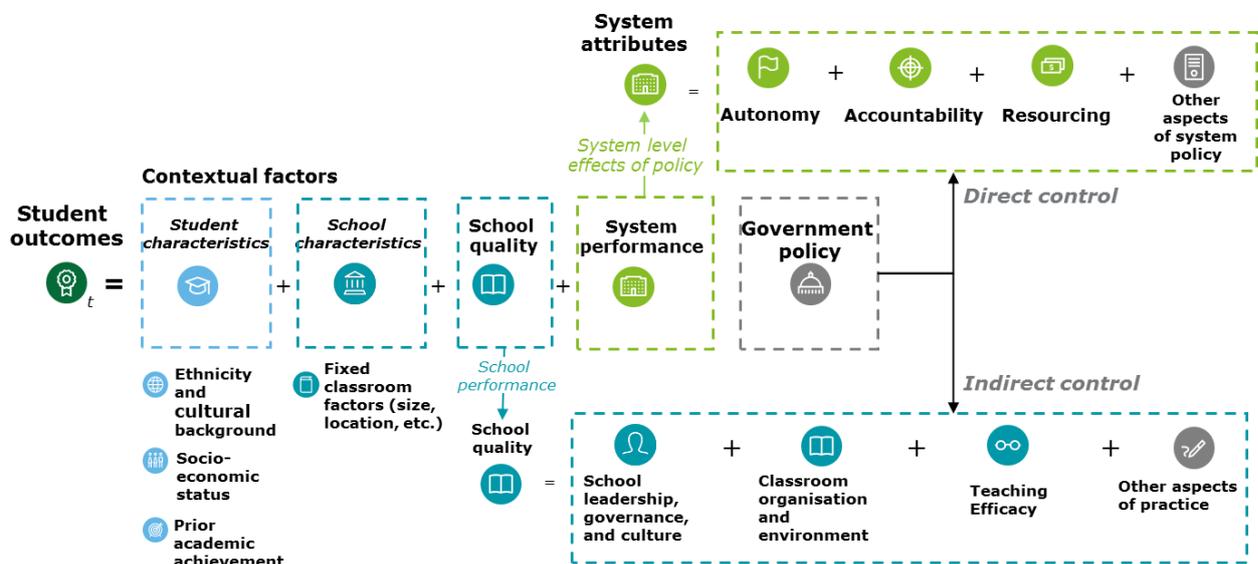
Multilevel contextual value added models have also utilised PISA data when investigating international differences in student performance. Most notably Fuchs and Woessmann (2004) use this approach to yield significant findings with the model accounting 85% of the between country variation, with roughly 25% of variation accruing to institutional variation.

2.3 A structural equation to isolate the drivers of school quality

The approach of this report builds on early work by Fuchs and Woessman (2004) to isolate measures of school quality and the factors which underpin it, with a specific focus on the Australian context.

This is achieved through the development of a structural equation of the drivers of student outcomes, and school quality, illustrated by Figure 2.4. This stylised equation builds on the measurement framework developed by this literature review. In this equation, contextual factors (the characteristics of students and schools that schools and government cannot change) combine with school quality and system performance (that is, the attributes of schools and school systems within the control of government) to determine outcomes. Chapter 3 of this report outlines how the PISA and TIMSS datasets may be used to develop detailed measures under each of these themes, and subsequently estimate effects of different drivers of student outcomes, and school quality.

Figure 2.4: Empirical approach to modelling drivers of student outcomes



Source: Deloitte Access Economics

3 Empirical methodology and data

This chapter explains the approach taken to matching the findings from the literature review to the PISA and TIMSS datasets, and how this is then used to test empirically for the core drivers of students' outcomes.

- Section 3.1 outlines the relevance of the PISA and TIMSS datasets from a statistical perspective.
- Section 3.2 details the 'value-add' econometric approach to analysing the relationship between student outcomes and the drivers of school quality.
- Section 3.3 considers the method used to aggregate individual question responses within the PISA and TIMSS databases into the themes of the measurement framework.
- Section 3.4 discusses PISA scores over time between Australian jurisdictions, and considers the feasibility of drawing a connection between system settings and outcomes.
- Section 3.5 outlines some of the limitations of the empirical analysis.

3.1 The PISA and TIMSS datasets

The primary datasets used in this empirical analysis are:

- The OECD Programme for International Student Assessment (PISA), which tests students' skills and learning progress in reading, mathematical and scientific literacy at age 15 (including students from years 9, 10 and 11).
- The IEA Trends in International Mathematics and Science Study (TIMSS), which tests students' curriculum knowledge and learning progress in maths and science in year 4 and year 8.¹⁰

Both datasets provide measures of student cognitive outcomes, from 2000 to 2015 in the case of PISA, and from 2003 to 2015 in the case of TIMSS. They also include details about:

- A range of student and school attributes (such as student socio-economic background, highest parental education, indigenous status and school location);
- Questions that measure students' engagement, wellbeing, and learning self-efficacy;
- Questions of school principals, teachers and students which capture a range of measures of approaches to school practice and management and the climate of the school (these were mapped to the themes within the measurement framework, identified through the literature review).

PISA and TIMSS take stratified random samples of students and schools in Australia, to ensure adequate representation among key characteristics at both the student and school level such as indigenous status and school sector. The approach to sampling in PISA changed between 2009 and 2012: the number of schools sampled more than doubled, with a commensurate reduction in the average number of students in each school. A summary of the number of students, schools and the average number of students in each school is provided in Table 3.1 below.

¹⁰ As a secondary dataset, the OECD Teaching and Learning International Survey (TALIS) may play a role in expanding on the analysis provided in this study. While it is not captured in this report, further analysis may incorporate TALIS into the empirical analysis, particularly where it is linked with PISA results (that is, through the 2012 TALIS-PISA study).

Table 3.1: Summary of student and schools in PISA and TIMSS data

	<i>PISA (15 year olds)</i>					
	2000	2003	2006	2009	2012	2015
Number of students	6,077	12,551	14,170	14,251	14,481	14,530
Number of schools	231	321	356	353	775	758
Average number of students	26.3	39.1	39.8	40.4	18.7	19.2
	<i>TIMSS (year 4)</i>					
	2003	2007	2011	2015		
Number of students	4,321	4,108	6,146	6,057		
Number of schools	204	229	280	287		
Average number of teachers/classes per school	1.4	1.1	1.3	1.2		
Average number of students per school	21.2	17.9	22.0	21.1		
Average number of students per class	19.5	13.0	14.0	12.2		
	<i>TIMSS (year 8)</i>					
	2003	2007	2011	2015		
Number of students	4,791	4,069	7,556	10,338		
Number of schools	207	228	277	285		
Number of teachers/classes	3.4	2.8	3.0	2.7		
Average number of students per school	23.1	17.8	27.3	36.3		
Average number of students per class	23.1	17.1	15.4	16.0		

Source: Deloitte Access Economics analysis of PISA and TIMSS data

As outlined in the table above, the PISA dataset captures a sample of students within a school, where these students may come from a range of different classrooms with different teachers. In contrast, the TIMSS dataset samples students from specific classes in a school (although in some schools, some students from outside the core class also sit the assessment). This means that it is possible to link teachers directly with students in the TIMSS dataset, but not for the PISA dataset.

The PISA and TIMSS datasets measure student outcomes through standardised tests, where estimates of achievement are based on statistical measures of the underlying academic ability of the student (that is, the tests of ability are measured with error, and the test results are interpreted as predictions of a student's academic ability). Scores are general presented in numerical terms and generally range from 360 to 630 for PISA and 390 to 610 for TIMSS.¹¹

Accounting for variations in age and year level

The PISA and TIMSS datasets include different measures of students learning achievement and measure students at different points in their schooling life. PISA measures students' skills in understanding and applying key concepts in the domains of Reading, Maths and Science; whereas TIMSS focuses on students' knowledge and learning progress of the Maths and Science curriculum. The timing of the PISA test also means that 15 year old students across multiple school year levels sit the test at the same time, whereas TIMSS captures students of different ages in years 4 and 8. The implications of this for this analysis include:

- Observations for Primary schools are only available through the TIMSS dataset; and

¹¹ This range is the 10th and 90th percentile of scores.

- Controls for school year level (and potentially age) are required when analysing the PISA data, whereas controls for student age are required when analysing the TIMSS data.

Because of the variation in schooling grades (years 9, 10 and 11) sampled in the PISA dataset, it is possible to derive a concordance between PISA scores and 'equivalent years of schooling'. For Australia, a difference in scores of around 30 points is estimated to be equivalent to one year of additional schooling. An equivalent measurement cannot be directly estimated for TIMSS, as all students who sit the two TIMSS tests (for year 4 and year 8 students) are in the same year level.

3.2 Understanding and measuring school quality

Following practice established in the literature, the contribution of school quality to student outcomes can be measured as the variation in student outcomes that remains after controlling for the observable characteristics of students and the school (Fuchs and Woessmann, 2004). In order to estimate school quality, so called 'value-added' models of student outcomes are used. In the context of PISA and TIMSS data, the value-added model of student outcomes is defined as:

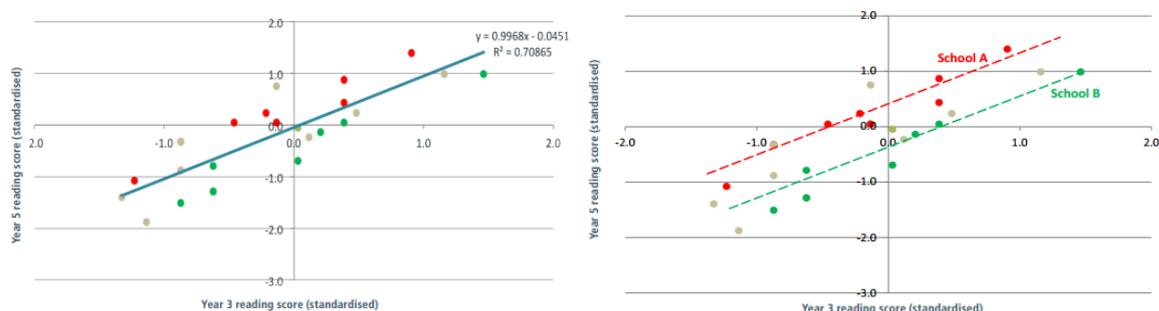
$$y_{ikjt} = X_i\beta + Z_k\Gamma + \gamma_{jt} + u_{kt} + \epsilon_i$$

Here, the score of student i in school k in setting¹² j at time t is modelled as a function of student characteristics $X_i\beta$, observable school characteristics $Z_k\Gamma$, and indicators for *specific policy settings* (such as differences across schooling systems) γ_{jt} . After controlling for these observable characteristics, what remains are idiosyncratic student level variations ϵ_i (driven by omitted variables that systematically influence student outcomes, such as innate ability or attitude; as well as randomness—such as falling sick on the test day) and school quality or *value-added* u_{kt} .

In plain language terms, value-added models (also known as multi-level models) are used to isolate the effects of differences in school practice on student performance, while controlling for observable student and school characteristics. Comparisons of performance between schools (in PISA) and across classrooms in schools (in TIMSS) can then be made on a 'like-for-like' basis to provide an estimate of the effect of school quality.

Chart 3.1 below provides a simple illustration of how school level value-added is estimated. In these charts, a variety of students with a variety of outcomes attend each school. After controlling for the starting ability of students, and assuming that each school has a similar profile of students, the vertical difference between the red and green lines can be considered as the additional value added by School A to the outcomes of its students, relative to School B. It is a measure of the additional score an average student receives simply by attending School A (due to its higher quality) instead of School B.

Chart 3.1: A simple illustration of school level value-added



Source: https://www.cese.nsw.gov.au/images/stories/PDF/VAPaper_v3-1Final.pdf

¹² Setting, here, refers to distinctions between *type* of school, for instance, by state (as different states set up their school systems in different ways) or by sector. These differences between settings are reflected in the 'indicators for specific policy settings' term γ_{jt} .

This approach measures the variation in outcomes resulting from the different practices and management of schools (in PISA) and across classrooms (in TIMSS).

3.3 Identifying and measuring the drivers of school and classroom quality

The second part of the analysis unpacks the relative importance of the various factors that make up school quality (u_{kt}) in the proposed model, to identify which ones are most strongly associated with student performance in Australia. Findings from this study can then be used to inform where governments might look to identify proven practices that will provide the strongest improvement to Australian schools in order to lift student outcomes.

Questions of school principals, teachers and students from PISA and TIMSS 2015 datasets were mapped to various aspects of school quality outlined in the measurement framework. These were included in the empirical model along with the student and school contextual factors to estimate student performance. This is achieved by replacing the indirect value-added measure of school quality with direct observations of school management and teaching practices.

The mapping process of representative questions for the themes and sub-themes in the measurement framework from 418 data items in PISA and 255 data items in TIMSS is explained in section 3.3.1 below. After incorporating direct observations of school quality using representative questions, the full student level model then becomes:

$$y_{ikjt} = X_i\beta + Z_k\Gamma + \gamma_{jt} + \theta_k C + u_{kt} + \epsilon_i$$

Here, the score of student i in school k in setting (all of Australia, metro, provincial or remote) j at time t is modelled explicitly as a function of the representative questions $\theta_k C$, along with student ($X_i\beta$) and school ($Z_k\Gamma$) characteristics, as well as differences across schooling systems γ_{jt} .

Including the drivers of school quality together (as opposed to the partial analysis undertaken in the previous section) allows us to isolate the effect of each sub-theme in driving student outcomes while controlling for both the contextual factors: in effect, controlling for the impact of the other themes. The standardised effect size of each representative question and theme can be calculated using the method described in Section 3.3.2 below.

Findings from the PISA dataset show the extent to which variation in student performance is driven by differences in the various aspects of *school* quality within the school and across schools. Findings from the TIMSS dataset, on the other hand, show the variation in aspects of school quality across *classrooms* of different schools. PISA captures information about teaching practice indirectly from the principal and through student perceptions of their teachers. PISA also provides more insight into the school climate while TIMSS has detailed information about the classroom and resourcing at the school. The TIMSS dataset offers direct insight into teaching practice and curriculum in the classroom from the teacher of students in the same class. Representative questions from PISA and TIMSS datasets would complement each other as each individual dataset has limitations in particular areas of the measurement framework.

Themes and sub-themes can then be ranked in order of relative importance, based on the extent to which their representative questions can explain variations in student performance. Themes that are considered more important aspects of school quality could identify substantial differences in school management and teaching practices across schools and classrooms that affect student outcomes in Australia. The relative importance of a theme is most commonly measured by two methods (Liu et al., 2014). A 'bottom-up' approach measures the difference in the coefficient of determination R^2 from adding all representative questions in a sub-theme, compared to just using the observed controls of student and school characteristics. Alternatively, the 'top-down' approach measures the difference in the R^2 measure that results from removing questions in a theme, compared to the full model (which includes the controls and all of the themes). This analysis takes the average of the two approaches to calculate the relative importance.

Robustness testing

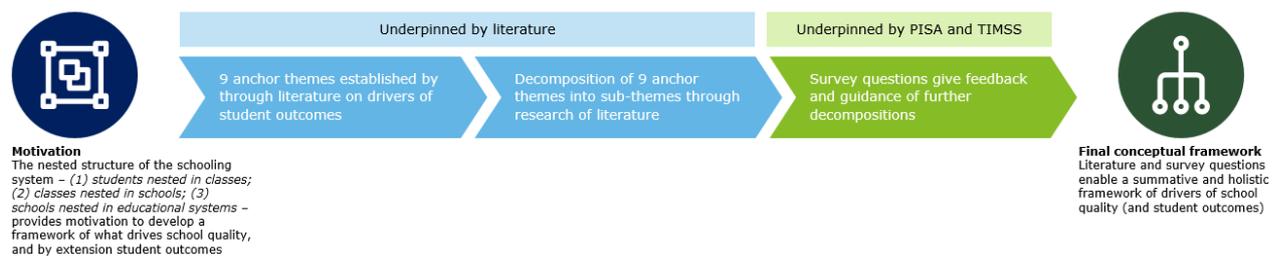
To test whether the identified set of representative questions capture the differences in school quality between schools, it is possible to rerun the multilevel regression with the representative questions and examine the residual school level value added that is yet to be accounted for.

It is also possible to undertake testing for different sub-samples, such as for different student and school characteristics, as well as across different time periods and jurisdictions. This allows for the identification of whether drivers of school quality differ between cohorts, and across time.

3.3.1 Mapping questions to themes and clustering

From a conceptual standpoint, the literature review outlined in Chapter 2 establishes a framework of the themes that drive student outcomes. Nine broad anchor themes capturing system, school and student level drivers of student outcomes guide the development of the measurement framework. The process of mapping PISA and TIMSS questions to these conceptual themes is achieved by decomposing these themes into a series of nested sub-themes, which capture more nuanced aspects of the drivers of student outcomes, broadly illustrated by Figure 3.1 below.

Figure 3.1: Expanding the measurement framework to encompass PISA and TIMSS questions



Source: Deloitte Access Economics

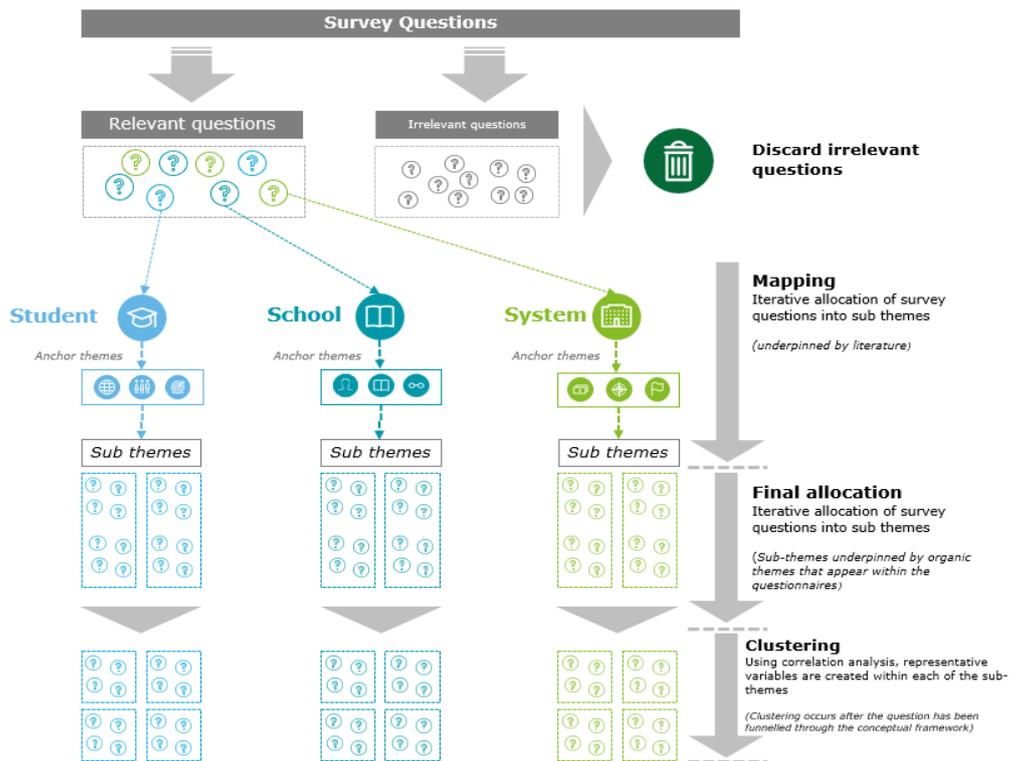
In approaching the mapping process, PISA and TIMSS questions are used to create instruments (that is, variables) that represent the themes identified in the literature, and which form the basis of the empirical modelling. That is, they comprise the independent variables which are modelled against student outcomes—the dependent variable. Given that the PISA and TIMSS questions generally relate to specific aspects of school practice and management within a given theme, a more precise definition of what each question is intending to capture is developed within the context of the overarching analytical framework.

The process of mapping survey questions to themes is summarised as follows (further illustrated in in Figure 3.2):

1. Each question is analysed and allocated to its 'anchor theme';
2. Questions that are unrelated or tenuously linked to the framework are discarded;
3. Repeat step 1 for questions in the anchor themes: analyse the question in each anchor theme and allocate it to the appropriate sub-theme; and
4. Continue until the survey question is allocated to the final level sub-theme.

It should be noted that this process generates a measurement framework that is more detailed than the broad outline motivated by the literature canvassed in Chapter 2. In particular, the detailed themes captured in the expanded measurement framework make use of the detail of the PISA and TIMSS datasets by establishing a set of detailed sub-themes within the broad categories previously established in Chapter 2.

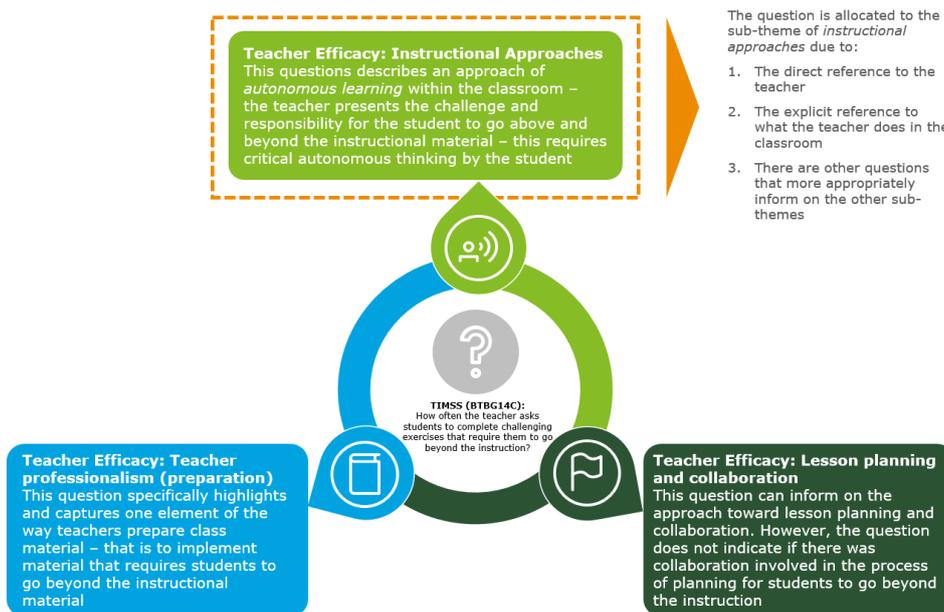
Figure 3.2: Process of mapping survey questions to measurement framework



Source: Deloitte Access Economics

Through this process, questions can be interpreted differently (as a result of containing different aspects of educational quality within the question itself) which makes the allocation of questions to sub-themes within the framework partly subjective. For example, the TIMSS question on how often the “teacher asks students to complete challenging exercises that require them to go beyond the instruction” (BTBG14C) can be interpreted in several ways, as seen in Figure 3.3.

Figure 3.3: Analysing the subjectivity of survey questions when mapping to themes



Source: Deloitte Access Economics

Within each of these sub-themes however, the wide array of survey questions invites the notion of further clustering questions together for the purpose of analysis.¹³ This is done to ascertain a representative question that is used in the estimation model, which is described in section 3.3.2.

Identifying system level factors in the PISA and TIMSS datasets

The PISA and TIMSS datasets include a broad range of questions and instruments which capture aspects of students' educational experience and school practice. These questions are necessarily answered at the level of the school, including teachers and students within the school. There are questions in these datasets **that capture aspects of the system context in which the school operates**. In this regard, they are not measures of school practice, but rather aspects of the system context over which educational authorities would generally be expected to exercise control. These questions broadly relate to the themes of autonomy, accountability and resourcing as established by the system, but from the perspective of the school.

These system focused questions are relatively few in number (particularly for TIMSS), and are **not direct observations of specific system policies** across Australia's schooling jurisdictions, and therefore do not lend themselves to direct assessments of the impact of any given policy. Rather instruments should be interpreted as instruments which capture aspects of system level attributes which may have implications for policy.

An alternative approach to considering the effects of system level attributes on student outcomes is to directly compare the performance of Australian schooling jurisdictions to differences in system levels policies, which relate directly to these themes. This more 'top-down' analysis of the impacts of system level drivers of student outcomes and school quality is considered in section 3.4 below.

¹³ Partial correlation analysis clusters variables together to derive a representative question (*not all sub-themes will necessarily contain multiple representative questions*)

An example of a complete mapping based on literature and survey questions can be seen through the anchor theme of *Teaching Efficacy*. In this framework, evidence from the literature has guided the structure to the 'level 3' sub-strata. Beyond this, the PISA and TIMSS survey questions were used to develop further decompositions of the drivers. This is presented in Figure 3.4 below. Detailed mappings for all nine anchor themes are included in Appendix A.

Example of mapping a question to a sub-theme

The below provides an example process of mapping a question to a sub-theme. This process accounts for the subjective nature of questions.

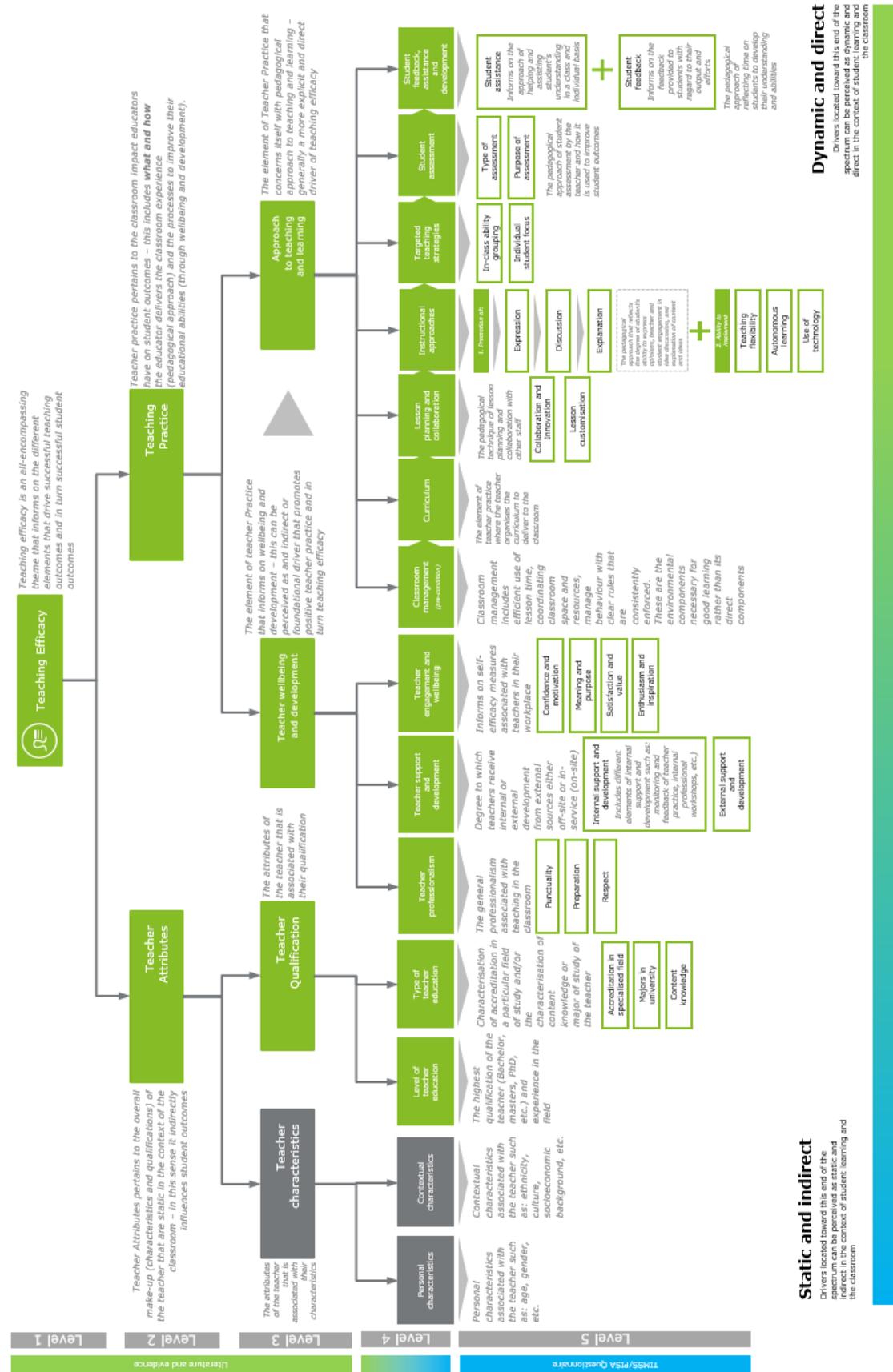
Question: BTBG14C: How often the teacher asks students to complete challenging exercises that require them to go beyond the instruction?



Source: TIMSS 2015 question

1. This question describes the notion of autonomous learning within the classroom environment, which is administered by the teacher. Applying the process of allocating the question to a sub-theme (*whilst accounting for the element of subjectivity*), the question strongly relates with the anchor theme of *Teaching Efficacy*.
2. Within *Teaching Efficacy*, the question is determined to be most related to *Teaching practice* (as it is also not an attribute of the Teacher).
3. Requiring students to complete challenging exercises beyond the instruction is an approach the educator takes to teaching and learning, in this sense, the questions is most suited to be allocated into the *Approach to teaching and learning* sub-theme
4. This question further pertains to a specific instructional approach, therefore it is further allocated to the *Instructional Approaches* sub-theme.
5. Finally, requiring the student to go beyond the instruction is strongly associated with autonomous learning techniques and results in the final location of the survey questionnaire.

Figure 3.4: Measurement framework mapping of *Teaching efficacy* theme derived from literature and detailed survey questions



Source: Deloitte Access Economics

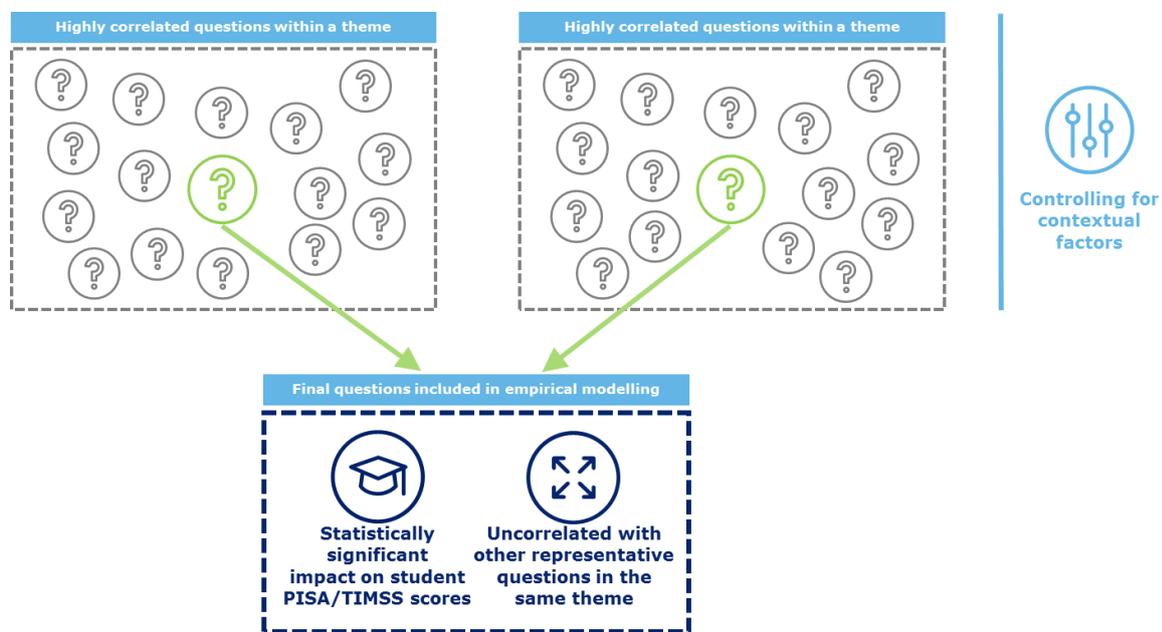
3.3.2 Identifying key questions for modelling and analysis

The top down approach of sorting survey questions into their appropriate themes can generate multiple questions that are associated with a final sub-theme, creating 'clusters' of questions for each sub-theme.

From a statistical standpoint, having multiple questions being modelled as individual independent variables increases the dimensionality of the model. Where these questions represent the same fundamental driver of student outcomes, estimation becomes less accurate, particularly where these variables are collinear in their effect.

In light of this, the methodology aims to reduce the number of questions, and the dimensionality issue, while capturing the underlying interpretation of the theme the questions are intended to represent. This is done through identifying one (or more) representative question(s) within each sub-theme, that are highly significant in explaining student outcomes, and that are correlated with the other questions in the sub-theme (that the chosen question effectively 'represents')—as illustrated in Figure 3.5 below. This procedure produces a set of final questions that are all significant predictors of student outcomes, while being largely uncorrelated with each other.

Figure 3.5: Illustration of the clustering method



Source: Deloitte Access Economics

A more detailed and technical description of this methodology is provided below.

Partial correlation analysis is used to determine the association between each relevant question and student outcome measures while controlling for student and school level contextual factors. To account for the degree of variability in responses to each question and allow for comparisons across questions, the standardised absolute effect and p-value is calculated using the partial correlation results. For categorical responses, this is given by the following equation:

$$b_i = \frac{\sum_j \sigma_{ij} |\beta_{ij}| p_{ij}}{\sum_j p_{ij}}$$

The standardised effect of question i (b_i) is given as the weighted average of the absolute standard effect of each response j relative to the omitted response (given as the standard deviation multiplied by the absolute value coefficient $\sigma_{ij} |\beta_{ij}|$). The weights are given by the likelihood of each response j being given relative to the omitted response (p_{ij}). A similar standardisation process is undertaken for the p-value.

For each sub-theme, the standardised p-values and the correlation between questions is used to cluster the full set of questions around a representative questions. The clustering process for each sub-theme is given as follows:

- Rank the questions in terms of their significance using the standardised p-value. Questions that do not meet a given significance threshold are discarded as not having a strong link to student performance.
- The question with the greatest significance is selected as the first representative question. Other questions within the sub-theme that are highly correlated with this question are 'represented' by this question, and are not included in the full analysis. The correlation threshold is set at 0.3, however this may be varied to provide a more or less strict threshold for inclusion, and varying the threshold can support robustness checks. The threshold was set at a relatively low level to reduce the full set of questions down to a tractable number. However, it is possible that questions that are weakly linked together have been included together. This parameter may be altered in future applications of the methodology
- If there are un-discarded questions after the first round of clustering, the question with the lowest correlation to the first representative question is chosen as the second representative question.
- The process is repeated until there are no more questions left (either unchosen as a representative question or un-discarded) in the sub-theme.
- The final list of representative questions are then included in the full model as drivers of student outcomes, representing the drivers of school quality.

It should be noted that this process is in part subjective. Due to the specific availability of information contained within PISA and TIMSS questionnaires, and further with the subjective mapping of questions, the results that arise from the proceeding analysis are predicated on the process outlined above and may be altered under alternative assumptions and qualitative heuristics. This is an inherent constraint to the analysis, however the analytical process utilised in this study is intuitive and aims to make best use with the data available.

3.4 Providing a link between system settings and jurisdictional performance

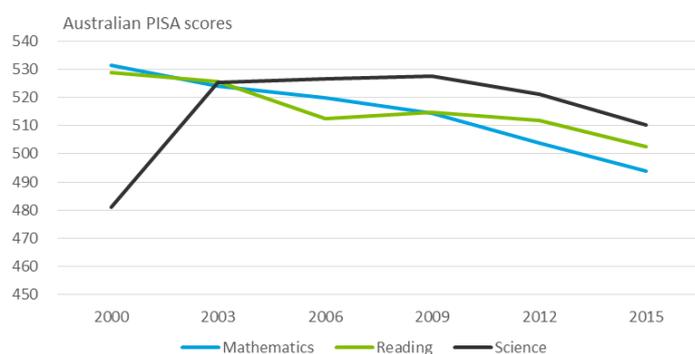
As outlined above, the PISA and TIMSS datasets do not generally include measures or instruments which directly capture the difference between system settings and policies in Australian jurisdictions, for the purposes of analysing how these aspects of policy may influence student outcomes in Australia.

Separate to the analysis outlined above, this analytical stream examines the longer term trends in student performance across different school systems in Australian jurisdictions, seeking to relate movements in these trends to differences in system settings at a thematic level. This is achieved by broadly mapping the system specific trends in performance estimated through the model outlined in section 3.3 above, against major historical system initiatives across Australian school jurisdictions and systems over the past 20 years. The following sections outline the approach to developing this evidence base and identifying the impact of these initiatives on student outcomes. The results of this analysis are presented in Appendix D of this report.

3.4.1 Revealing systemic changes in performance

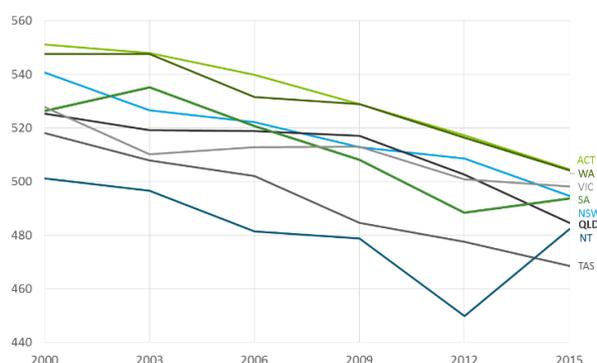
Chart 3.2 and Chart 3.3 below show a summary of the recent trends in PISA scores for Australian jurisdictions, from 2000-2015. While there is a clear downward trend in system level performance (measured in terms of average student outcomes) over this period, there is also an **observable difference in the permanent levels of performance across jurisdictions**, and the level of change that has occurred over time. The ACT and Western Australia, in absolute terms, have been the highest performers on PISA, while Tasmania and the Northern Territory's scores have been consistently lower than those of the other states. Although all states have declined over the period, the Northern Territory has significantly improved between 2012 and 2015 (noting the limitations of the small sample sizes in that jurisdiction); Victoria has had the smallest decline among jurisdictions while Tasmania has had the largest decline.

Chart 3.2: Trends in PISA scores for Australia, 2000-2015



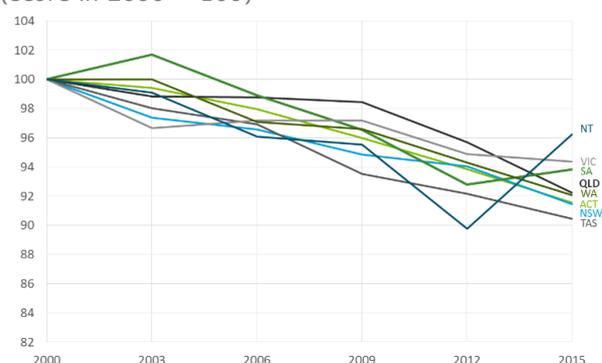
Source: Deloitte Access Economics analysis of PISA data

Chart 3.3: Trends in PISA Mathematics scores by state, 2000-2015



Source: Deloitte Access Economics analysis of PISA data

Chart 3.4: Trends in PISA Mathematics scores by state, 2000-2015 (score in 2000 = 100)



The results above do not account for the fact that each jurisdiction has a different—and changing—cohort of students across each test year. As identified in the literature review, both school and student contextual factors, such as socio-economic status, can have a significant impact on student results. When measuring the performance of a jurisdiction in contributing to student outcomes, it is necessary to control for these contextual factors to isolate the impact of each jurisdiction’s schooling system at a high level.

Following the method of controlling for observable contextual factors, estimated system level effects—all else being equal—represent differences in the system settings of jurisdictions within Australia. In practice, it is possible that these estimated system level effects will be relatively small, given much of the variation in student outcomes can be explained by student and school level factors. Where significant variations in system level effects are identified, both across jurisdictions and over time, these may be linked to observed changes in governments’ approach to schooling that has occurred in these jurisdictions.

3.4.2 Desktop review of schooling system settings and initiatives

To unpack what system-level differences are contributing to variation between jurisdictions, a desktop review of ‘grey literature’ has identified the most significant schooling system changes in Australia over the past 15-20 years.¹⁴

¹⁴ This research included Commonwealth, State and Territory education department policy and procedure documents, policy reviews from the peer-reviewed literature and bodies such as the Australian Centre for Educational Research (ACER) and policy instruments such as industrial agreements.

- The number of government initiatives in the schooling sector across Australia is practically innumerable. DET, for this project, has provided a preliminary list of 329 schooling reforms. Building on the evidence from the literature review with respect to the system level factors that drive educational outcomes, a **set of themes** was developed that can be used to tractability categorise and prioritise the nature of differences and reforms across Australian jurisdictions. Using the framework from the literature review on drivers of school quality, hypotheses can be formed about the pathways through which system changes ultimately affect student outcomes. Some initiatives may have an indirect impact on outcomes because they involve changing high-level system settings. Others which aim to more directly shape school practice may have a more direct impact on outcomes.
- Across each of the themes, the next steps of the approach were to identify evidence of **permanent and timing differences** in the approaches of government, across Australian schooling jurisdictions and over time. This is required to properly parametrise the empirical analysis, and form effective hypotheses on how certain changes may influence student outcomes.
 - **Permanent differences** are **identifiable differences between schooling systems that are fixed** (at least for some period of time) **between jurisdictions**, such that observable differences in system performance may be associated with these observed differences in the approaches of government.
 - **Timing differences** are where a **comparable initiative is implemented at different times across different jurisdictions**, so that changes in outcomes over time can be associated with the implementation of certain initiatives.

In practice, some aspects of schooling systems across jurisdictions may have both permanent and timing differences.

- Based on the quality of available evidence, and the level at which initiatives are implemented, hypotheses were drawn about the **expected impact of future changes**.
 - Where possible, system-level settings will be mapped against their expected effectiveness based on factors in the literature review.
 - Those initiatives which rely on school level implementation are harder to assess the impact of from PISA results, as PISA itself relies on an unidentified sample of schools. Accordingly, it is hard to determine which school-level interventions have been implemented in the PISA dataset. For these, case studies will be used to examine initiatives, to illustrate the ways in which government can change school practice.

3.5 Limitations to the analysis

As with any original empirical study, the approach used in this study is not without its limitations. In particular:

- One of the central contributions of this research involves the mapping of PISA and TIMSS questions to conceptual themes and drivers of student outcomes, to identify and estimate the relative importance of these different factors. A key limitation of this work lies in the quality of the instruments available for this study (that is, the usefulness of the variables in the PISA and TIMSS datasets), as well as the accuracy and appropriateness of the mapping exercise.
 - Ultimately, the mapping process identified above is subjective, and there will exist limitations in the appropriateness of the interpretation of different questions when representing different conceptual themes.
 - The results of this analysis should be interpreted in the context of the particular PISA and TIMSS questions which are used in the modelling, which may have vague or inconclusive links to aspects of practice.
 - The approach to this study has been to develop a methodology which will extract the most meaningful insights available from this dataset, and future research which builds on this methodology would be expected to refine and enhance the mapping process and interpretation of the results accordingly.
- The strength of the conclusions made in this report rely on the quality of the evidence available in the PISA and TIMSS datasets, and the way in which the data is used to represent the measurement framework developed through this study. While these datasets are extensive, they are by no means comprehensive. In some cases, where the evidence shows that a given

theme is more or less important in explaining student outcomes, this may result from the fact that no effective instrument was available to demonstrate its impact.

- While the PISA and TIMSS tests and surveys are conducted every 3-4 years, they are not longitudinal in nature. That is, the students and schools sampled for the test are not common across years. This means that it is not possible to capture dynamic effects of practice on student outcomes over time and that the analysis can only be conducted with contemporaneous observations of school practice and student performance.
 - In this sense, it is difficult to determine whether the effects of certain factors on students' outcomes are causal in nature. Unlike a random controlled trial (RCT) this study uses naturally occurring variations in the data to estimate the effects of different factors.
- Following from the static nature of the PISA and TIMSS tests, it is not generally possible to account for the prior achievement of students when measuring the effect of practice on their outcomes. Because this is a 'point-in-time' association of school and teacher quality with outcomes, the effects of aspects of school quality on student outcomes may be over-stated relative to their actual effect (for example, in the presence of ability based school selection policies, which may attribute more of a student's outcomes to a school's practices and less to a student's prior ability) or under-stated relative to their actual effects (as prior achievement also contains information about the accumulated effects of school quality on the student).
- While this study briefly considers aspects of student outcomes beyond academic achievement on tests (for example, student engagement and wellbeing) they do not form a significant component of the analysis. This is not to downplay the importance of these other measures of outcomes in understanding the drivers of student outcomes and school quality in Australia.
- Similarly, the regression models which this utilises focus on mean student outcomes, and do not consider the effects on the distribution of outcomes (or the typical outcomes for student with certain characteristics). Further extensions to this research may look to explore these other effects in more detail, due to their implications for understanding how drivers of school quality can influence different aspects of student performance in varying contexts.

4 Findings from the analysis

This Chapter outlines the findings from the empirical analysis, in line with the methodology set out in Chapter 3.

- Section 4.1 presents results from the first stage of the analysis—a 'value-added' modelling approach to estimating school and classroom quality;
- Sections 4.2-4.5 outline the findings from the analysis of the different drivers of school quality, based on the variables identified through the mapping process set out in section 3.3.

The results from the analysis set out in section 3.4 are provided in Appendix D of this report.

4.1 Measuring school quality and value-added

Multi-level modelling utilises the nested nature of students within schools to isolate the effects of differences in school practice on student performance from the attributes of individual students, such as their socio-economic status, age and other contextual factors. As the isolated effect can be considered the 'value-added' by the school (over and above the other contextual drivers of outcomes), they are also often referred to as value-added models (see Section 2.2 of this report).

Multi-level models are important in education research as just looking at the absolute performance of schools does not give an accurate reflection of the school's quality. This is because students are not randomly placed within schools, but instead students tend to congregate, based on factors such as socio-economic status and prior academic achievement, which themselves have an effect on school performance outside the actions of the school.

By controlling for observed student and school characteristics, comparisons of performance between schools are made on a 'like-for-like' basis. The results can be interpreted as the effect of moving the same student from one school to another. Note that this is a relative measure that compares against the value added by the average school. Consequently, the results are meaningful in understanding the variation between schools rather than the absolute value added by schools.

Table 4.1 outlines the contextual factors that have been controlled for in the multi-level modelling for PISA. They cover both contextual factors at the student level and at the school and jurisdictional level. A similar set of controls have been used for the TIMSS dataset, however it is smaller due to limitations in the questionnaire. This can be found in Appendix C: Table C.1.

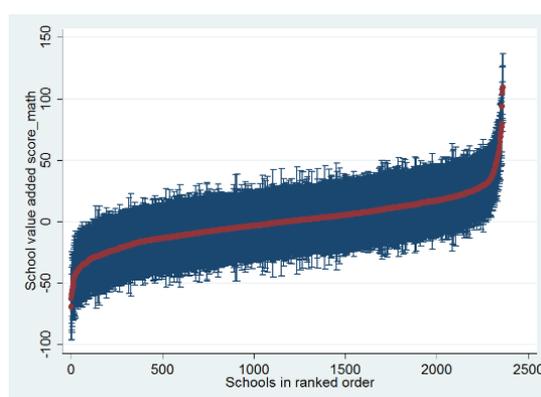
It is notable that prior academic achievement is not observable through the PISA and TIMSS datasets, which is a limitation of the analysis, to the extent that non-random assignment of students to classes (based on ability) is not properly controlled for in the modelling. It should be noted though that the analysis does explicitly control for whether the school is academically selective or not.

Table 4.1: Student and school characteristic controls, PISA

Student characteristics	School characteristics	Other characteristics
ESCS ¹⁵	School ESCS ¹⁶	State ¹⁷
School year level ¹⁸	School proportion of migrants ¹⁹	
Migrant status (including first and second generation migrants)	Total enrolment	
Age (months from 15 years old)	Student-teacher ratio	
Parental occupation ²⁰	Government/non-government status ²¹	
Gender	School location ²²	
Indigenous status ²³	Regional or local education authority approves student admission	
Number of books at home		

The distribution in estimated school value added to students' maths scores in PISA is shown in Chart 4.1. The data combines all schools from 2000 to 2015. The majority of schools are estimated to have 'value-added' performance which is statistically indistinguishable from each other, as given by the concentration of school value added estimates centred on zero. Indeed, only 15% of schools have value added estimates significantly different from the average school.

Chart 4.1: Distribution in school value added to student PISA maths scores



Source: Deloitte Access Economics analysis of PISA data

¹⁵ The PISA index of economic, social and cultural status (ESCS) uses a number of indices to consider students' socio-economic status. The particular set of indicators used has varied between test years to include some combination of highest educational or occupational status of parents; family wealth; cultural possessions; and home educational resources.

¹⁶ Captured separately to account for the potential impact of concentration of disadvantage.

¹⁷ State data is not available for 2000 data.

¹⁸ Students taking PISA are 15 years old. 75% of students taking PISA nationally in 2015 were in Year 10. Depending on the age at which they started school, and if they have repeated a year, they may be in a different year of schooling.

¹⁹ Captured separately to account for the potential impact of concentration of migrants.

²⁰ Because ESCS is an index of a number of factors, this study separately controls for parental occupation to consider any effect it has alone. For instance, a parent may be in a highly qualified field, but still face lower earnings than other people with a similar qualification. The student's low ESCS score would not reflect their parents' occupation and education level, which may still make a positive contribution to the student's outcome.

²¹ The government/non-government status of schools is not available for all years or schools. For those schools with missing data, the status has been imputed based on the mix of funding the school receives, the number of enrolments, and the student-teacher ratio.

²² School location status is not available for 2000 data.

²³ Indigenous status not available for 2000 data.

However, the difference between the top and bottom 10% of schools in terms of value added still represents significant differences for individual students, at a difference of 44 points, which is equivalent to approximately 1.5 schooling years.²⁴ This suggests that if Australia were able to lift the performance of a school from the bottom 10% to the top 10%, it would be equivalent to 1.5 years of schooling for those students. Indeed, the estimated value-added of schools to student outcomes is jointly significant in predicting student outcomes across the Australian schooling system.

These results should not be interpreted as school quality having only a small impact on student outcomes, but rather as an indication of the variation in quality that exists across schools. Under this methodology, a high performing system could be one where all schools contribute equally to the outcomes of students and hence where there is no measurable variation in value-added made by schools. Indeed, this form of value-added analysis says nothing about the overall level of quality of Australian schools, only variation in quality.

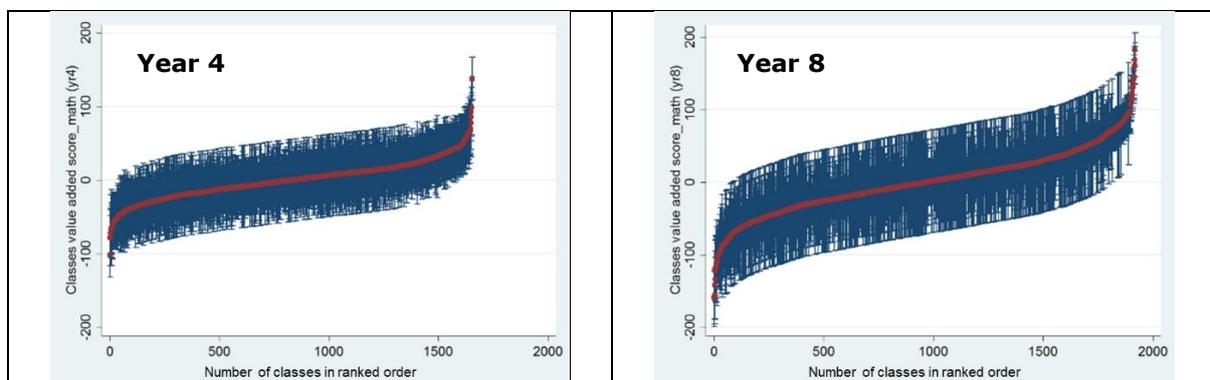
The results above isolate the school level value added while controlling for differences in jurisdictions (that is, Australian states and territories). However, the results can be disaggregated to both show the average jurisdiction's value added compared to the national average, as well as differences in the distribution of value added estimates across jurisdictions. This jurisdictional analysis is further explored in Appendix D of this report, when considering the link between performance and policy at a jurisdictional level in Australia.

Similar results can be produced for the TIMSS datasets. Recognising that the TIMSS dataset nests students within classes with a single teacher, it is possible to estimate the 'value-added' provided by classrooms to student outcomes, which can be interpreted as a measure of classroom quality. The contribution of classroom quality to student maths scores in TIMSS for year 4 and year 8 are given in Chart 4.2.

Comparing the TIMSS results to the PISA results, the contribution made by classroom quality to student outcomes is typically higher than the contribution made by school quality alone. This is particularly the case for year 8, where 37% of classrooms have a value added estimate statistically different from zero. The equivalent figure is 20% for year 4 classrooms.

This finding is consistent with evidence from similar studies which emphasise the significance of individual teaching practice and the classroom environment in driving student outcomes, irrespective of the specific school environment. The difference in value added between the top and bottom 10% of classrooms is 106 points for year 8 and 65 points for year 4. While it is not possible to convert this score into equivalent years of schooling, this contribution is highly significant, particularly when compared to the relatively modest contribution made by value-added measures of school quality.

Chart 4.2: Distribution in classroom value added to student TIMSS maths scores



Source: Deloitte Access Economics analysis of PISA and TIMSS data

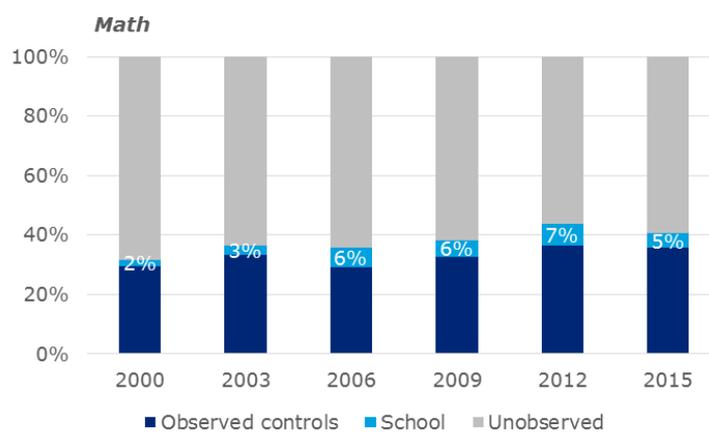
²⁴ Deloitte Access Economics makes the simplifying assumption that 30 points in PISA scores is approximately equivalent to one schooling year (ACER, 2016).

The results of the multi-level modelling can be disaggregated to show the proportion of the variation in student outcomes that is explained by different contextual factors at the student and school level, relative to the value-add contribution of schools and teachers.

The results from the analysis show that approximately 30% of the variation in PISA scores is accounted for by observable student and school contextual characteristics, including student and school level Economic Social and Cultural Status (ESCS), the location of school, the system (that is, government or non-government) and the jurisdiction. Then, variations in 'school quality' explain between 2% and 7% of the total variation in student outcomes observed across the country (Chart 4.3). This represents approximately an additional 0.3 to 1.2 years of schooling. The share of value added explained by schools have broadly remained constant over time. It increased from 2% in 2000 to 7% in 2012 before decreasing to 5% in 2015.

These results are lower than those presented in the 2016 report by Deloitte Access Economics, which found that schools accounted for 6% of variation in maths scores in 2003 and 14% of maths scores in 2012.²⁵ The difference is due to the inclusion of additional controls for state, age and grade in the current modelling. This is because these factors are not random at the school level. For instance, the grade of students will cluster due to differences in starting school ages between states. With the addition of these controls, the variation explained by them are now attributed to the observed controls. Nonetheless, a significant proportion of variation remains unobserved and therefore cannot be attributed to particular causes. This includes, but is not limited to students' ability, prior achievement, self-efficacy and motivation, and their home learning environment.

Chart 4.3: Accounting for variation in student PISA maths scores over time, 2000-2015



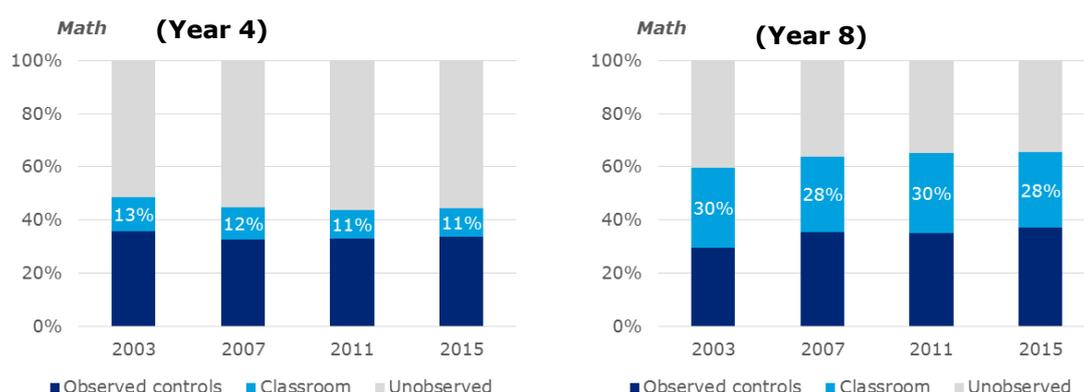
Source: Deloitte Access Economics analysis of PISA data

The TIMSS data shows that between 11%-13% of the variation in year 4 maths scores is due to differences at the classroom level. This increases to 28% to 30% for year 8 students. This is higher than the school level results from PISA, and suggests that differences in classroom quality is a more significant factor in driving student outcomes, as discussed above.

Compared to the PISA results, there has been less variation in the contribution of classroom quality over time. In very broad terms, these results demonstrate that a modest increase in the variation of school quality (but not classroom quality) in Australia has coincided with an overall decline in performance. This indicates that variation in school quality may be associated with more unequal distributions of quality teachers (and effective teaching) across the schooling system, disproportionately affecting certain students and schools who are potentially falling behind as others succeed in improving their performance and outcomes.

²⁵ See Deloitte Access Economics (2016), *The economic impact of improving school quality*, Chapter 8.

Chart 4.4: Accounting for variation in student TIMSS maths scores over time, 2000-2015



Source: Deloitte Access Economics analysis of TIMSS data

In addition to the role of schools in explaining variation in student academic performance (as proxied by their scores in the TIMSS and PISA test), it is also possible to isolate the school (and classroom) effects on non-academic outcome measures, including student sense of belonging, and engagement. The results for TIMSS for 2015 are presented in Table 4.2 below.

Table 4.2: Classroom effects on non-academic outcome measures, 2015

Student outcome measure	Proportion of student level variation explained by teacher effects	
	Year 4	Year 8
Student sense of school belonging	8%	10%
Student wellbeing ²⁶	4%	5%
Like learning maths	5%	9%
Engaged with teaching in maths	9%	19%
Confidence in maths	3%	7%
Maths score	11%	28%

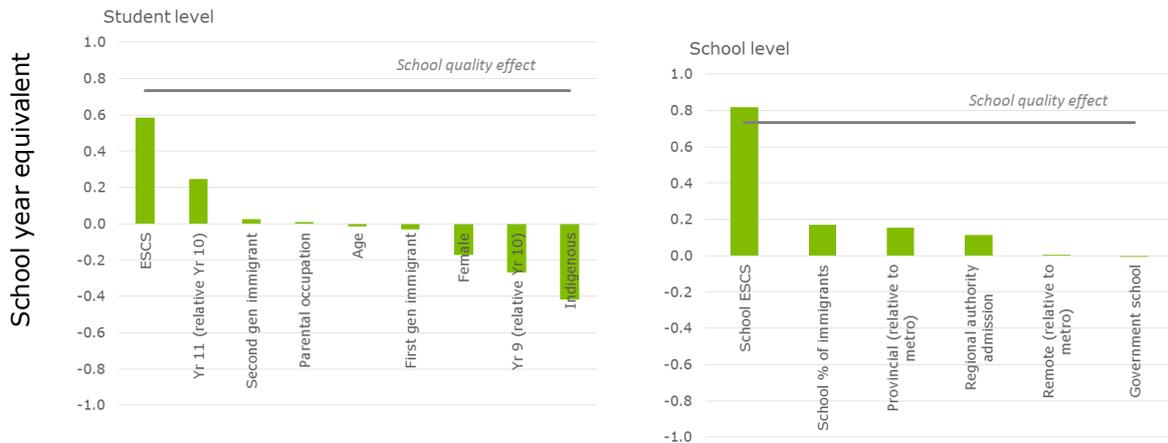
Source: Deloitte Access Economics analysis of TIMSS data

To place the contribution of teacher and school quality to student outcomes in context, it is possible to compare the effect of school quality against the standardised (that is, the typical) effects of different contextual characteristics at both the student and school level. The results for PISA are presented in Chart 4.5, while the results for TIMSS are presented in Chart 4.6.

These results emphasise the importance of socio-economic status (measured using the ESCS index) and Indigenous status in predicting student learning outcomes. This typical effect of school quality is estimated to be lower than the typical effect of school level socio-economic status on student outcomes, which indicates that typical variations in school quality do not offset the typical impact of factors of educational disadvantage. The typical effects of other contextual characteristics are generally modest—for example, the effect of being in a government school (relative to a non-government school)—while positive—is very small.

²⁶ Measured based on student reports of bullying

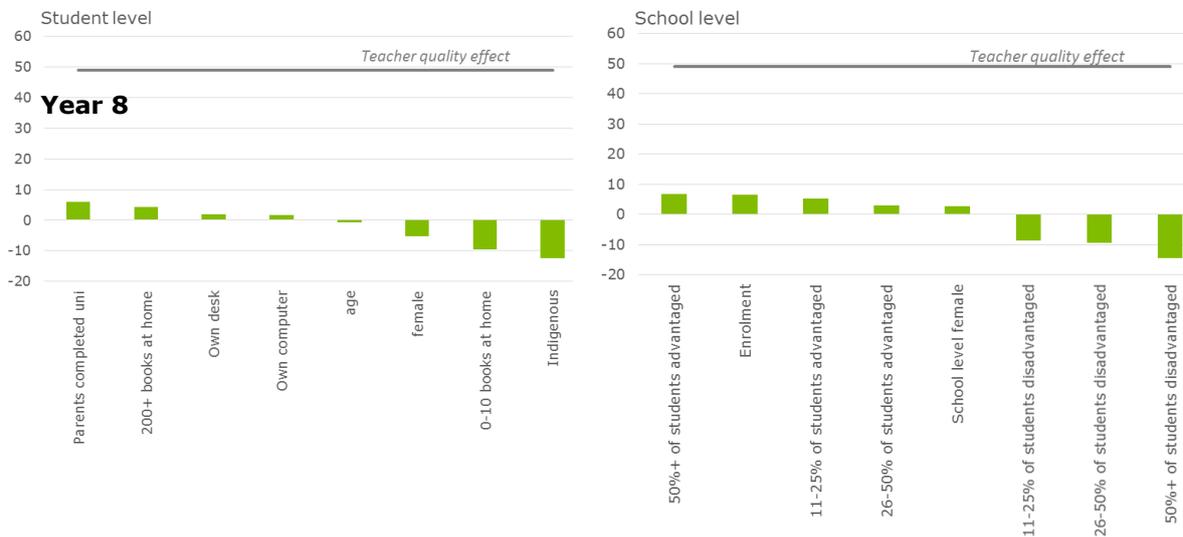
Chart 4.5: The relative effects of factors which influence student outcomes (PISA maths), 2015



Source: Deloitte Access Economics analysis of PISA data

The TIMSS data suggests that while school quality alone is unable to overcome contextual disadvantage, it can when combined with classroom quality. Variations in classroom quality typically contributes 50 points to student outcomes, compared to -20 to 10 points for contextual factors. While the effects at the student and school levels are likely to be underestimated in the TIMSS data due to their correlated nature—for instance parental education and home possessions, and the proportion of economic advantage and disadvantage within a school—classroom quality is still likely able to offset their typical impact on student outcomes.

Chart 4.6: The relative effects of factors which influence student outcomes (TIMSS maths), 2015



Source: Deloitte Access Economics analysis of TIMSS data

School quality and economic outcomes

Based on previous research by Deloitte Access Economics on the *Economic contribution of improving school quality*, it is possible to consider the economic effects of increases in school quality, estimated through the above modelling and analysis.

Assuming an initiative was able to lift the performance of the bottom 10% of schools to the 10% of schools. This would add 1.5 years of schooling to those students, and is equivalent to 44 points. Given an average of 494 points in PISA 2015 maths scores, this would raise the average PISA score by 4.4 points, or 0.9% overall.

Given a 0.9% increase to average PISA scores, GDP will increase by 0.14% once the effect is fully realised in the labour force. This is equivalent to a gain of \$2.0 billion (if the effects were fully realised in 2017).

As classroom quality is a larger contributor to student outcomes compared to school equality (see TIMSS results), it is likely that improving classroom quality will have a larger positive impact on the economy.

The factors that contribute to school and teaching efficacy are further explored and estimated in the following sections, providing insights into how government may change their approach to achieve such improvements in quality.

4.2 Identifying and measuring the drivers of school quality

As a part of the PISA and TIMSS tests, questionnaires for students, school principals and teachers (TIMSS only) are administered to gain insights into the aspects of school practice and management that are associated with student achievement (among other objectives). In PISA, there are over 400 individual questions related to teaching practice and school conditions. For TIMSS, there are over 250 questions. Given the large number of questions, it is neither tractable nor desirable to work with the full set of questions.

Sections 3.3.1 and 3.3.2 outline the methodology used to categorise, cluster and prioritise these questions for inclusion in the modelling of student outcomes, and school quality. This is achieved by selecting a set of representative questions for analysis, which are statistically significant predictors of student outcomes, and which are correlated with a range of other (ontologically similar) questions (which the questions subsequently 'represent').

For the questions identified under the most granular strata of the measurement framework set out in section 3.3.1, a p-value significance threshold of 0.15 is used to reduce the set of questions to relevant and significant predictors of student outcomes. This excludes questions that are not significant in explaining variation in student outcomes after controlling for contextual factors. Then, the questions are ranked in terms of their p-value significance, and a correlation threshold of 0.3 is used to reduce the set further. This removes questions that move in line with one another and are consequently 'represented' by the most significant identified question in a given theme.

An illustrative example of the process is seen for the theme 'instructional approaches'. Of the 14 PISA questions considered under this theme, four are excluded for not having a strong relationship with student PISA scores. The remaining eight questions (which include how often the teacher demonstrates an idea, how often the teacher discusses students' questions etc.), are all highly correlated with each other. Consequently, the set is reduced to one key question, *how often students are allowed to design their own experiments (in science class)*, which had the lowest p-value among the correlated set of questions. This question, while specific in nature, is used to represent the general theme of 'instructional approaches' and—by virtue of the established methodology—will approximately provide a measure of the effects of the other, omitted, questions under the 'instructional approaches' theme.

A similar process is repeated across each of the sub-themes constructed in the measurement framework. Overall, the 418 questions of the PISA dataset are reduced to 63 representative questions that represent 22 sub-themes relating to quality at the school and system level. A similar process is undertaken for TIMSS, reducing the set of 255 questions to 76 questions representing 24 sub-themes.

The set of representative questions is then added to model with the student and contextual controls to understand the relative importance of each driver of school quality, as outlined in

section 3.3.3 above. A full list of the representative questions mapped under each theme has been provided as an accompanying addendum to this report.

Figure 4.1: Flow chart for selecting representative questions



Source: Deloitte Access Economics

Representative questions for 'system level factors' in the PISA and TIMSS datasets

The process of identifying only significant representative questions based on an analysis of the relationship between questions and student outcomes eliminates a range of questions from the full model. In particular, this process eliminates all questions considered to be related to the system level theme of accountability. Further, the model includes relatively few questions that relate to aspects of autonomy in PISA, and no such questions in TIMSS.

In this regard, this 'bottom-up' approach to analysing the drivers of student outcomes and school quality is limited in its ability to highlight the contribution of these more enabling system factors like accountability and autonomy. Further, these questions are not instruments for specific policies that may relate to these aspects of Australian education systems. The relationship between more direct observations of system settings and system level performance is explored further outside of this core model, in section 4.3 of this report.

4.3 Drivers of school quality and their relative importance

This section tests the relative importance of each school quality 'driver' in explaining variation in student PISA or TIMSS scores, following the method set out in section 3.3.3. In simple terms, a driver that explains more variation in student outcomes, and has a greater standardised effect on student outcomes, is considered more important.

The results from this analysis at the sub-theme level (level 2 in the overarching measurement framework) are presented below in Table 4.3. These results provide an estimate of how much each theme contributes to differences, or variation, in student outcomes. This is presented as a percentage of the total variation in student outcomes.

For instance, variation in *teaching practice* explains the largest variation in student scores, at 6.1% for PISA maths scores, and 13.1% for TIMSS year 8 (and 3.9% of TIMSS year 4) math scores. When all the drivers of school quality are added, 13.8% of variation in PISA maths scores can be explained. Similarly, 27.5% and 8.4% of the variation in TIMSS math scores can be explained respectively at the year 8 and year 4 levels.

The variation in outcomes explained by school quality drivers in the PISA dataset is higher than the variation explained by the value-added measures outlined in section 4.1. This is likely the case as the PISA questions capture aspects of practice at the student level, which may vary within schools, and therefore capture unobserved effects of varying classroom quality. The variation in

outcomes explained by school quality drivers in the TIMSS dataset is broadly in line with the value-added measures set out in section 4.1. The effects of observed aspects of school and teaching practice in TIMSS explain slightly less variation in student outcomes than the value-added measure of classroom quality, which suggests that some important aspects of teaching practice (which influence student outcomes) remain unobserved in this analysis.

Across the data sets, teaching practice is consistently found as the most important theme driving student outcomes. Classroom environment; school leadership, governance and culture; and material based resourcing are also ranked among the most important drivers in all three of the test data sets.

Factors relating to system autonomy—as identified by schools—are found to be less important in predicting student outcomes than other, more direct, aspects of teaching practice. This is to be expected, as factors relating to autonomy vary at a system level, and would be expected to influence student outcomes and school quality through their impact on aspects of school practice. Given this analysis considers the contribution of different factors of student outcomes at a school level in a ‘bottom-up’ fashion, it does not provide a direct link to aspects of policy which relate to factors such as autonomy, and the link between these policies and initiatives on system level performance. Section 4.3 of this report considers the performance of Australia’s schooling jurisdictions in the context of these more direct differences in policy initiatives, while Chapter 5 of this report makes some observations regarding the implications of these findings for policy-makers in Australia.

Table 4.3: Relative importance of drivers of school quality themes

Themes	PISA (maths)	TIMSS (maths, year 8)	TIMSS (maths, year 4)
All school drivers	13.8%	27.5%	8.4%
Teaching efficacy - Practice	6.1%	13.1%	3.9%
Classroom organisation and environment - Environment	4.8%	3.9%	0.9%
School leadership, governance and culture*	2.1%	4.5%	2.0%
Resourcing - Material based	0.5%	2.2%	0.7%
Autonomy ²⁷	0.4%	n/a	n/a
Classroom organisation and environment - Organisation	0.2%	3.6%	1.1%
Teaching efficacy - Attributes	0.1%	1.4%	0.3%
Accountability	0.0%	n/a	n/a
Resourcing - Curriculum and staff based	n/a	0.3%	0.2%

The relative importance of each driver in the analysis of each data set is indicated by the colour of the cell: a darker blue indicates that a driver was of greater relative importance, while lighter blue shading indicates that a driver was of lower relative importance.

* School leadership, governance and culture is an anchor theme.

Source: Deloitte Access Economics analysis of PISA and TIMSS data

The following sections disaggregate results for the most significant identified themes into further detail.

It should be noted that the methodology developed for this study generates an extremely detailed set of results and findings, of which the below is only a subset. The questions which comprise this analysis are often complex and nuanced in nature, and this report does not seek to examine and evaluate the implications which arise from all of the findings which have been generated. Rather, it intends to consider the implications at a more strategic and summative level. While some specific and detailed commentary is provided in the following sections, a more comprehensive analysis and

²⁷ The ‘system level’ theme of autonomy is found to be less important, largely because there are relatively few questions that can effectively capture aspects relating to this theme in the PISA and TIMSS datasets. Similarly, there are limited PISA and TIMSS questions that can act as instruments for the theme of ‘accountability’ at the system level, which mean it is not included in the analysis. This limitation of the data should be noted when interpreting the relative importance of the measures outlined here.

discussion of implications for policy which arise from all of this study's empirical findings remains an area for further work.

Teaching efficacy – teaching practice

'Teaching practice' pertains to the professional attributes and attitudes of teachers, as well as their approaches to delivering teaching and learning in the classroom. This covers many elements of practice, including what the educator delivers to the class (curriculum), how the educator delivers to the class (instructional approaches), how the educator tracks learning outcomes (student assessment) and whether the educator provides one on one assistance as required (targeted teaching strategies). Teacher wellbeing and development factors (engagement and wellbeing, professionalism) are more indirect components of teaching practice that influence the effectiveness of teaching.

The relative importance of the sub-themes within teaching practice are presented below. 'Instructional approaches' covers variations in the methods used by teachers to run their classes: the pedagogical approaches that teacher employ, the extent to which students may express opinions, teacher and student engagement in idea generation and discussion, and explanation of content and ideas. This sub-theme has the largest explanatory power for high school students, explaining 2.9% of the variation in student PISA maths scores, 4.3% of year 8 TIMSS maths scores. By way of comparison, the relative contribution made by school level 'value-add' in explaining student outcomes is around 5%. This suggests that differences in teaching practice between schools are the most significant driver of the variation of outcomes explained by schools.

In contrast, instructional approaches are relatively less important in explaining variation in TIMSS year 4 maths scores after controlling for contextual factors, accounting for 0.9% of the variation in scores. 'Instructional approaches' as a driver of quality is represented by 9 key questions at the year 8 level representing variations in approaches to teaching, including *how often the teacher asks students to work on challenging problems or problems for which there is no immediately obvious method of solution*. At the year 4 level, it is represented by 4 questions, including *whether the teacher asks students to work on problems while the teacher is occupied by other tasks*, and *how often students are allowed to design or plan experiments or investigations*. Two questions – *how often students are permitted to use calculators in class*, and *how often students work in mixed abilities groups* – are common to both grades.

In the PISA dataset, professionalism (exhibiting the conduct and behaviours expected of a teacher) is also a relatively important sub-theme, explaining 3.1% of the variation in student maths scores. In contrast, it is relatively less important for TIMSS, explaining just 0.20% of the variation in maths scores for year 8 students. This, in part, may be the result of the questions available for mapping in each test.²⁸ For the PISA questionnaires, the theme is represented by two key questions: (1) *whether teachers disciplined the student more harshly compared to other students*, and (2) *whether the teacher called on the student less often than other students*. In contrast, professionalism in the TIMSS dataset does not directly focus on their behaviour towards students, and is represented by the question *whether teacher arriving late is a problem for the school*. The PISA dataset also suggests that the employment of targeted teaching strategies contribute to significant variation, and that in particular, student achievement is greater where teachers who are more likely to provide individual help when an individual student has difficulties.

In TIMSS, the most important sub-theme for explaining variation in year 8 maths results is 'instructional approaches'. This is followed by teacher *engagement and wellbeing* (the extent to which teachers are motivated and prepared to teach). This is represented by questions relating to whether *they feel prepared to teacher certain topics in mathematics*. This explains 4.0% of the variation in maths scores for year 8 students, and 1.2% of the variation in maths scores for year 4 students. Engagement and wellbeing is relatively more important for explaining variation on year 4 student outcomes (compared to teaching practice overall).

²⁸ Further information about the broader limitations of the mapping process are outlined in section 3.3.1.

For year 4 students, curriculum (that is, the type of content covered by the teacher) is also more important in explaining variation relative to year 8 students. This could be because at the lower grades, there are fundamental numeracy topics that should be covered. Consequently, ensuring what students learn (whether through the curriculum or through teacher confidence in particular topics) is more important than variation in instructional approaches themselves. The proportion of variation explained by the sub-themes within teaching practice is given in Table 4.4.

Table 4.4: Relative importance of themes – teaching practice

Themes (Teaching efficacy - practice)	PISA (maths)	TIMSS (maths, year 8)	TIMSS (maths, year 4)
Approach to teaching and learning - Assessment	0.09%	1.94%	0.02%
Approach to teaching and learning - Instructional approaches	2.85%	4.28%	0.92%
Approach to teaching and learning - Curriculum	n/a	2.83%	1.18%
Approach to teaching and learning - Lesson planning and collaboration	n/a	1.03%	0.25%
Approach to teaching and learning - Targeted teaching strategies	2.00%	n/a	n/a
Wellbeing and development - Professionalism	3.06%	0.20%	0.46%
Wellbeing and development - Engagement and wellbeing	n/a	4.02%	1.18%
Wellbeing and development - Support and development	0.08%	0.16%	0.07%
Total teaching efficacy - practice	6.07%	13.11%	3.88%

Source: Deloitte Access Economics analysis of TIMSS data

It is also possible to directly interpret how the representative questions affect student PISA scores both directionally and in terms of the effect size. The partial effects in Table 4.5 show the relationship when only controlling for student and context, and the *full* effects show the relationship when controlling for student and context, as well as other teaching practice and school quality drivers. The theme 'instructional approaches' is represented by the question *how often students are allowed to design their own experiments* in PISA 2015 (Table 4.5).

For instance, the table shows that holding contextual factors and the effects of other drivers of school quality fixed, moving a student to a class that is a standard deviation more likely to allow students to design their own experiments compared to the average class will lower their PISA maths scores by 2.45 points. Note that, because PISA 2015 had a science focus, it did not directly survey some aspects of maths teaching. Accordingly, there is no direct link between the science-based instructional approach question and student maths outcomes.

Table 4.5: Significant representative questions – teaching practice (PISA)

	Question	Standardised effect	
		Full	Partial
Support and development	Teachers in our school cooperate by exchanging ideas or material when teaching specific units or series of lessons.	5.56**	1.52
Professionalism	Teachers are more likely than average to discipline me more harshly than other students.	-3.8*	-9.53
Professionalism	The teacher is more likely than average to say something insulting to me in front of others.	-2.85*	-9.17
Professionalism	Teachers more likely to than average call on me less often than they called on other students.	-2.75*	-6.79
Instructional approaches	Students are more often than average allowed to design their own experiments (science class).	-2.45**	-8.47
Assessment	Teacher developed tests are used in school to award certification to students.	2.33*	1.59

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of TIMSS data

In TIMSS, more frequent use of instructional approaches that focus on certain tasks compared to the average teacher has a positive effect on student TIMSS results in maths. This includes assigning specific tasks such as *working on problems as a whole class with guidance from the teacher*, and *allowing students to use calculators*.

Furthermore, encouraging students to work in mixed abilities groups more often than average has a positive relationship with student PISA scores. This emphasises the value of collaboration, as also seen through the positive effect when students are also encouraged to work on problems in class.

This finding supports previous research around the use of ability-based groupings, and raises the importance of understanding the difference between *within-class* targeted teaching and *between-class* ability grouping. Targeted teaching strategies—where teachers of one class use different approaches to educating students with different levels of prior ability in a topic area—play a significant role in determining outcomes. However, grouping students into entirely separate classes based on their ability has been found to have particularly negative equity impacts on ‘low-track’ groups (as discussed in section 2.1.2 above).

It should be noted that some of the results presented here are not intuitive in terms of their effect size. In some cases, controlling for other factors can make interpreting results difficult. For example, the question: “Teacher work together with other teachers to try out new ideas more often than average” is found to have a negative effect on student outcomes. Controlling for other measures of practice though, this question may be identifying those less experienced teachers, who seek support to develop lesson plans and improve their practice. In general, it should also be noted that these questions are intended to be representative of a broader range of correlated questions which sit within each practice theme, and any direct interpretation of the focus of a particular question should be made with caution.

Table 4.6: Significant representative questions – teaching practice (TIMSS, year 8)

		Standardised effect	
	Question	Full	Partial
Instructional approaches	Teacher asks students to work on problems as a whole class with teacher guidance less often than average	-9.99***	4.73
Lesson planning and collaboration	Teacher work together with other teachers to try out new ideas more often than average	-8.76***	-9.55
Support and development	Teacher has participated in professional development for addressing individual students’ needs in past two years	8.39***	-3.03
Instructional approaches	Teacher asks students to work in same ability groups more often than average	-8.34**	5.5
Engagement and wellbeing	Teacher finds work full of meaning and purpose less often than average	-7.53***	-8.01
Engagement and wellbeing	Teacher feels safe at school less than average	6.8**	-15.04
Curriculum	Teacher taught algebra simplification last year (compared to teaching it this year)	-5.95***	11.3
Engagement and wellbeing	Teacher less prepared to teach comparing and ordering rational numbers compared to average	-5.37***	-8.82
Assessment	Teacher places less emphasis on classroom tests compared to average	5.22**	-6.92
Curriculum	Teacher taught Cartesian planes last year (compared to teaching it this year)	-5.09**	9.99
Assessment	Teacher asks students to correct their own homework less often than average	-4.81**	-13.98
Curriculum	Teacher taught properties of functions last year (compared to not yet teaching it)	-4.28**	-12.06

	Question	Standardised effect	
		Full	Partial
Curriculum	Teacher taught simplifying and evaluating algebraic expressions last year (compared to teaching it this year)	3.76*	8
Curriculum	Teacher taught characteristics of data sets last year (compared to teaching it this year)	3.67*	6.03
Instructional approaches	Teacher allows students to use calculators during lesson less often than average	-3.43*	-7.43
Assessment	Teacher places less emphasis on assessment of students' ongoing work than average	-2.85*	7.97
Curriculum	Teacher yet to teach using appropriate measurement formulas (compared to teaching it this year)	-2.53*	-8.71
Curriculum	Teacher taught computing with rational numbers this year (compared to teaching it last year)	-1.86*	-13.74

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of TIMSS data

Overall, these results emphasise the importance of key aspects of teaching practice, such as targeted teaching and effective instructional approaches. Notably, these pedagogically focused themes are more important than those themes which relate to curriculum and assessment, and the process of lesson planning and collaboration.

It is not within the scope of this study to examine the findings for each question and identify implications for practice and policy. Such analysis and evaluation of implications may be conducted by building on the methodology and evidence base established through this study.

Classroom organisation and environment

Classroom environment, as a sub-theme, covers the learning climate of a classroom that students participate in, separate from the explicit efforts of the teacher. Class order and cohesion (which describes the level of discipline and student behaviour in class) is represented by key questions such as *how often students do not listen to what the teacher says (PISA)*, *the degree to which teaching is hindered by uninterested students (TIMSS year 8)*, and *whether the students behave in an orderly manner (year 4)*. Peer culture (which describes the manner that students interact with each other) is represented in PISA by *whether students enjoy working in teams*, and *whether students enjoy considering different perspectives*. The proportion of variation explained by the sub-themes within teaching practice is given in Table 4.4. Overall, classroom environment is found to be important for explaining variation in student scores, particularly at the high school level.

Classroom organisation, on the other hand, describes the approach used by a school in determining the number of students in each class (compositional strategies), the support offered to teachers (in the form of teaching aides) (teacher working conditions), and the method used by schools in grouping students into subject classes (practice driven organisation). Each of these were found to be of relatively little importance in explaining variation in student outcomes, with the exception of compositional strategies in TIMSS (year 8). In this case, a greater number of students in a class was associated with stronger student outcomes. This result should be interpreted with caution: with the significant number of contextual controls in the model, class size may contain other information about the school that are positively associated with student outcomes. Student gender mix and grade mix are also aspects of composition strategy but were not measured in the datasets.

Table 4.7: Relative importance of themes – classroom organisation and environment (percentage of variation in student outcomes explained)

Themes (Classroom organisation and environment - environment)	PISA (maths)	TIMSS (maths, year 8)	TIMSS (maths, year 4)
Environment - Class order and cohesion	2.17%	3.93%	0.86%
Environment - Peer culture	3.16%	n/a	n/a
Total classroom organisation and environment - Environment	4.85%	3.93%	0.86%
Organisation - Compositional strategies of classroom organisation	n/a	2.64%	0.35%
Organisation - Practice driven organisation	0.09%	0.15%	0.19%
Organisation - Teacher working conditions	0.10%	0.62%	0.53%
Total classroom organisation and environment - Organisation	0.19%	3.60%	1.11%

Source: Deloitte Access Economics analysis of TIMSS data

Based on the TIMSS year 8 analysis, moving a student from a class with an average level of uninterested students to one that is a standard deviation more interested would raise the student's score in maths by 6.28 points. This suggests that increased classroom segregation, such as removing disinterested students to separate education environments, is likely to have a disproportionately negative impact the outcomes of residualised students and classrooms. This is because there will be fewer students in the class to prevent teaching being hindered by uninterested students/encourage different perspectives, and promote overall order and cohesion in the classroom.

Notwithstanding the observations made above regarding teaching strategies that encourage within class ability groupings of students, this analysis suggests that grouping students based on ability to different classes could have a positive effect after controlling for other factors (such as classroom order and cohesion, like student interest levels). It should be noted that this finding is significant at the 10 percent level, where the other key findings in this sub-theme are significant at the 1 percent level. This suggests that this is an area where further exploration and research may be required.

This is particularly true because of the complex impact that ability streaming may have on particular types of students. For example, the results presented here represent an average impact of the effects of ability-based 'streaming' on student outcomes, and do not consider the separate impacts of students with low or high levels of current achievement. As has been noted previously, ability-based streaming can have disproportionately negative effects on those students who are streamed into lower achievement levels classes.

Table 4.8: Significant representative questions – classroom organisation and environment (TIMSS, year 8)

	Question	Standardised effect	
		Full	Partial
Compositional strategies of classroom organisation	Number of students in class	12.09***	17.79
Order and cohesion	Students behave in an orderly manner less often than average	-9.25***	-20.4
Teacher working conditions	Number of students experience difficulties understanding spoken English in class	-6.35***	-5.36
Order and cohesion	Teaching is less limited than average by uninterested students	6.28***	18.06
Practice driven organisation	Student achievement is used to assign students to maths classes	5.06*	3.87
Order and cohesion	Teacher's teaching is limited by students with physical disabilities more than average	-1.98**	-14.02

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of TIMSS data

School leadership, governance and culture

School leadership, governance and culture covers a number of sub-themes on the overall school mission and specific policies, as well as the degree of parental and staff involvement with the governance of the school. It functions as an effective pre-condition that implicitly enables improved student outcomes to be achieved via the conditions in which teachers work and deliver classes. This includes the following:

- **School policy, mission and goals** informs the high level strategic direction and objectives set out by the school. It is represented by key questions including whether *the school uses data to plan specific action for school development*, whether *improvement exists at the school*, and whether *the school offers assistance with schoolwork*.
- **Principal attributes, culture and integration** describes the degree to which principals are involved with the governance and culture of the school. It is represented by key questions including the *highest level of education obtained by the principal*, whether *the school implemented any measures in teacher development*, and whether the principal *promotes teaching practices based on recent educational research*.
- **Parental culture and involvement** describes the degree to which parents are involved with the governance and culture of the school. It is represented by questions including whether *the school provides a welcoming and accepting atmosphere for parents to get involved*, and the degree of *parental commitment to ensure that students are ready to learn*.
- **Staff culture and involvement** describes the degree to which teaching staff are involved with the governance and culture of the school. It is represented by questions including the principal's opinion on *teacher's ability to inspire students*, and the frequency that principals *engage teachers to help build a school culture of continuous improvement*.

The relative importance of the sub-themes in explaining variation in student maths scores is summarised in Table 4.9. Using the PISA dataset, principal attributes, culture and integration is the most important sub-theme, but is relatively less important for TIMSS. This is due in part to the different ways in which the questions are framed. Whereas the PISA questionnaire focuses on the *role* of principals, in encouraging teacher professional development, and driving school improvements, the TIMSS questionnaire focuses on the *attribute* of principals, such as their years of experience. In contrast, the TIMSS dataset finds that policy, mission and goals is the most important sub-theme within school leadership.

Table 4.9: Relative importance of themes – school leadership, governance and culture (percentage of variation in student outcomes explained)

Themes (School leadership, governance and culture)	PISA (maths)	TIMSS (maths, year 8)	TIMSS (maths, year 4)
School leadership, governance and culture - Policy, mission and goals	0.33%	2.18%	1.11%
School leadership, governance and culture - Parental culture and involvement	0.34%	1.07%	0.61%
School leadership, governance and culture - Principal attributes, culture and integration	1.42%	1.14%	0.05%
School leadership, governance and culture - Staff culture and involvement	0.21%	0.67%	0.45%
Total school leadership, governance and culture	2.07%	4.55%	1.96%

Source: Deloitte Access Economics analysis of PISA and TIMSS data

It is difficult to compare the findings of this report with those of Deloitte Access Economics' 2016 report, which similarly tried to show how different aspects of schooling explained variation in PISA maths scores (in 2003 and 2013). This is due to differences in the measurement framework used for identifying and measuring the drivers of school quality. For instance, the 2016 analysis found that the amount of homework received by students, the disciplinary climate in class, and classroom management were the aspects of schooling that drove the greatest variation in PISA test scores. In contrast, teacher morale, teacher behaviour and the quality of material resources were not found to be significant across either the 2003 or 2012 cohorts.

The disciplinary climate in class maps directly to classroom environment (class order and cohesion), while questions on the amount of homework received by students is mapped to teaching practice (assessment). These are both found to be relatively important as drivers of school quality. Given the different approach and focus of the previous report, the impact of instructional approaches was not explicitly examined. Further analysis of previous years of data from PISA and TIMSS may provide further insights on the relative importance of these different factors.

4.4 Relative importance of school quality drivers across outcome measures

Student academic achievement is not the only measure of student outcomes. It is also possible to disaggregate the relative importance of different school quality drivers on a range of other measures of student outcomes. These outcomes can be both important in and of themselves (such as outcomes for student safety and wellbeing), and also be indirectly linked to student academic achievement. Within the TIMSS dataset, aggregate student outcome variables have been created based on student responses to individual questions. They include:

- **Student sense of school belonging** covers ten questions including whether '*students enjoy school*', '*whether they feel like they belong in the school*', '*whether teachers are fair to them*', and '*whether they are proud to go to the school*'.
- **Student safety and wellbeing (bullying)** covers nine questions on how often they experienced bullying behaviour. It covers questions such as '*how often they were made fun of by other students*', and '*how often they were left out of other students' games or activities*'.
- **Student engagement with teaching** covers ten questions for the students including '*I know what my teacher expects me to do*', '*my teacher does a variety of things to help us learn*', '*my teacher has clear answers to my questions*', and '*my teacher gives us interesting things to do*'.

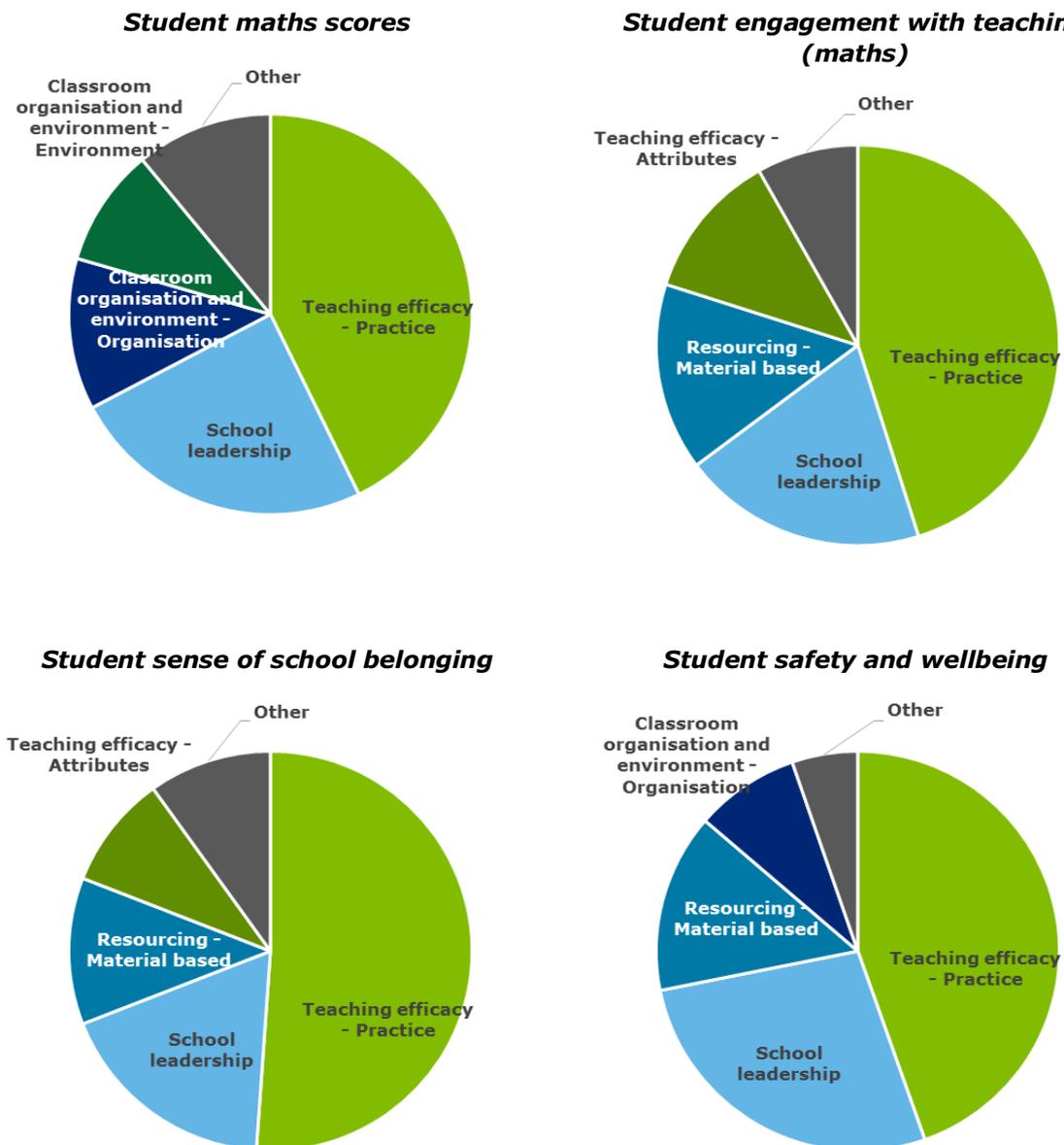
Chart 4.7 below highlights the top four most important themes in explaining variation in each of the student outcomes. Teaching practice is consistently the most important theme across all outcome measures. More than student maths scores, it explains a larger share of variation in student safety and wellbeing, and student sense of school belonging.

Teacher attributes are generally not found to be important in determining student outcomes, with the exception of their estimated contribution to student engagement with teaching. Material based resourcing is also found to be important for student engagement with teaching. It is possible that

attributes such as experience with teaching allow teachers to explain concepts more clearly, while better resources help to add variety to teaching methods, which would encourage student engagement.

Classroom environment is found to be relatively unimportant in explaining student safety and wellbeing, with classroom organisation the more relevant theme.

Chart 4.7: Relative importance of quality drivers by student outcome measure (TIMSS, year 8, 2015)



Source: Deloitte Access Economics analysis of PISA and TIMSS data

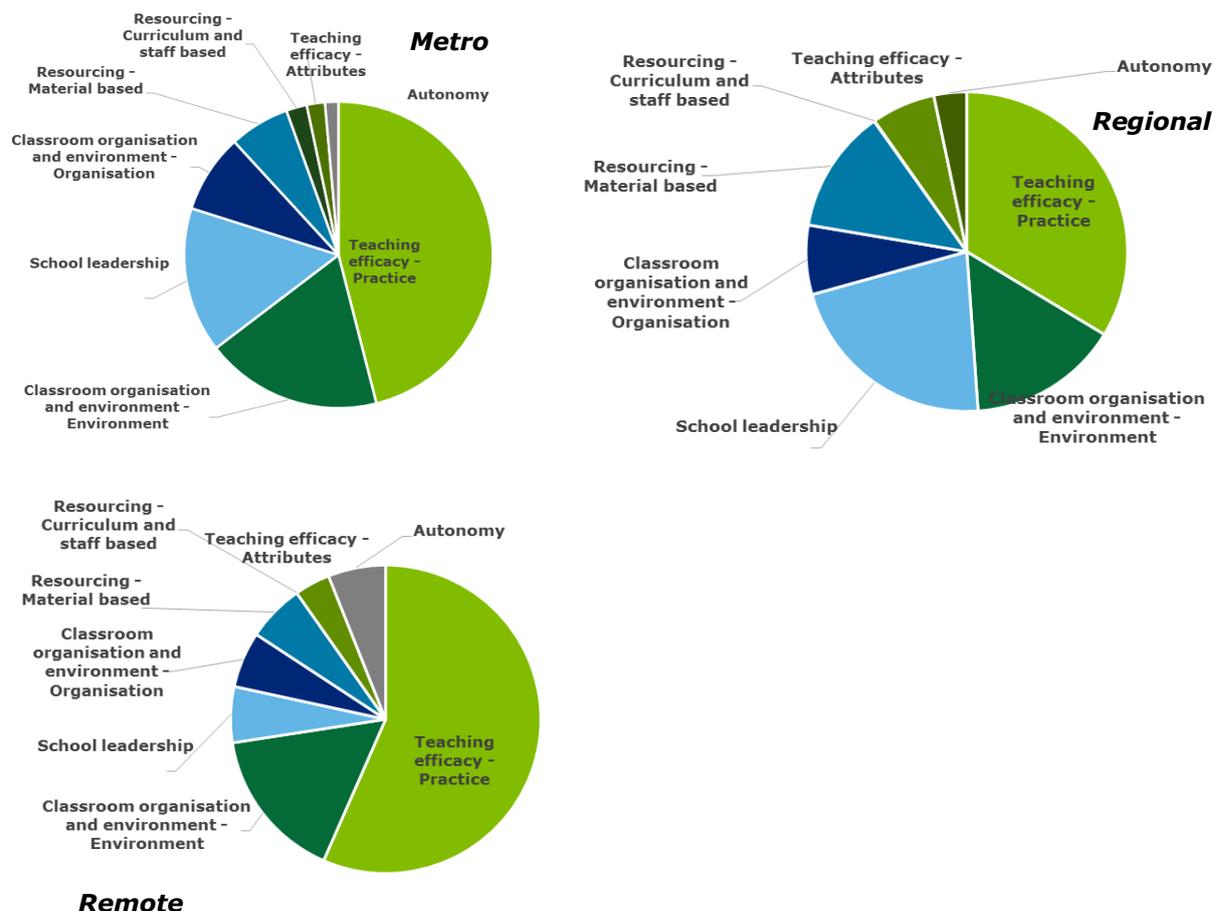
4.5 Relative importance of school quality drivers across geographies

The relative importance of the themes in explaining variation in student outcomes also differs between geographies, that is, whether a school is located within a metropolitan area, a provincial area, or remote area.

Chart 4.8 shows the relative importance of school quality drivers averaged for PISA and TIMSS (maths, year 8). Teaching practice is the most important driver for all geographies. School leadership and classroom organisation are relatively more important for schools located within metropolitan and provincial areas. In contrast, classroom environment and material based resourcing is relatively more important for schools located in remote areas.

School leadership and classroom organisation are potentially more important for metropolitan and regional schools given the student population is more diverse. Consequently, these schools may require a more targeted approach to school leadership and classroom organisation to meet to the needs of the students. In contrast, these factors are relatively less important for non-metropolitan schools as the student population is likely to be more homogenous. Detailed findings for the geographic analysis is given in Appendix C.

Chart 4.8: Relative importance of quality drivers by geography (2015)



Source: Deloitte Access Economics analysis of PISA and TIMSS data

In summary, across these data sets, **teaching practice and the classroom environment are consistently found to be the most important theme driving student outcomes**. This result aligns with the literature; however, the magnitude of the importance of teaching practice relative to other factors, such as school leadership and autonomy, is notable. These results demonstrate that, as a driver of student outcomes, teaching practice (such as approaches to teaching and learning) is significantly more important than teacher attributes (such as teacher qualifications).

These findings have implications for this report’s understanding of the role that the strategic direction of schooling systems play in driving improvements in student outcomes across Australia’s education system. This is discussed in the concluding Chapter 5 of this report, and further evidenced with respect to variations in historical system level policies and practices in Appendix D.

5 Implications for government and future research directions

The empirical analysis above builds upon the existing literature to demonstrate where the most significant gains can be made in improving students' outcomes in Australian schools. The analysis—in part—confirms what is already known, that differences in teaching practice, and classroom organisation and environment, explain the most significant amounts of variation between schools and classrooms. This provides confidence and assurance for policymakers in targeting policies towards the most important drivers of school quality, based on a comprehensive, and perhaps most significantly, Australia-specific evidence base.

This section draws out the results of the empirical analysis of most importance to policy makers, examines the current policy landscape, and considers how the empirical results of this work can be implemented in policy, before concluding with a summary of the directions for future research building on this work.

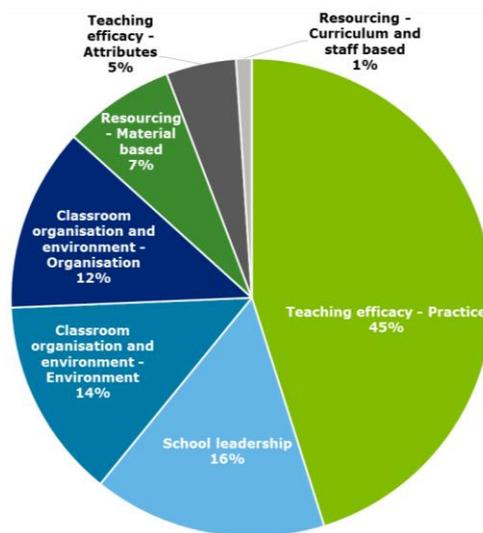
5.1 An evidence base on schooling systems, practice and performance in Australia

5.1.1 Results of empirical analysis of drivers of school quality

Variations in average 'school quality'— the differences in the practice and management of schools that affect student outcomes— explain around 5% of the total variation in student outcomes in Australia. The contribution made by variations in classroom quality is typically much higher than the contribution made by school quality: as much as 28% of the variation on student outcomes is explained by variations in what happens in the classroom. Digging deeper into the drivers of this variation in school and classroom factors, this analysis finds that variations in **teaching practice** have the most significant effect on student outcomes (Chart 5.1):

- Variations in teaching practice explain 6.1% of the variation in PISA maths scores, and 13.1% of the variation in TIMSS math scores (for year 8 students).
- The **effect of teaching practice** is estimated to be **twice as significant** as the next most significant driver of school quality.
- Within teaching practice, the **most significant drivers of student outcomes** include instructional approaches, targeted teaching strategies, professional approaches to teaching and learning, and strategies for student engagement and wellbeing. These drivers emphasise the importance of teachers adapting their approaches to meet the needs of individual students in classrooms of varying levels of interest and ability.
- The second largest contribution to student outcomes is the **classroom environment**, which explains up to 7.5% of the variation in student outcomes.
- In relative terms, teaching practice and the classroom environment (including the engagement and wellbeing of students) account for more than half of the variation in student outcomes attributed to the identified 'drivers of school quality'.
- Other relatively important factors include school leadership and material based resourcing.
- The contributions of other factors are generally smaller. For example:
 - Measures of school autonomy at the school level identified through PISA contribute very little to the variation in student outcomes after controlling for other factors; and
 - Differences in teacher attributes (including qualifications) are also found to be less important in explaining variations in student outcomes, relative to teaching practice.

Chart 5.1: Illustration of the overall relative importance of quality drivers (TIMSS Year 8 maths scores)



Source: Deloitte Access Economics analysis of TIMSS data

5.1.1 A framework for understanding the role of government in the school system

For policy purposes, knowing relative importance/relative contributions is a significant improvement from just knowing rankings of different practice-based drivers from the literature. In particular, it is more instructive when it comes to decisions about where to invest additional resources in schools. This study demonstrates that the key drivers of outcomes improvement – teaching practice, classroom environment and school leadership – are held at a school level. What, then, leads to practical changes in teaching practice and in classrooms that results in improved student outcomes?

The link between these drivers and the role of government is not always clear, as school practice and management can be influenced and guided by a range of sources, across different systems, jurisdictions, regions and local communities.

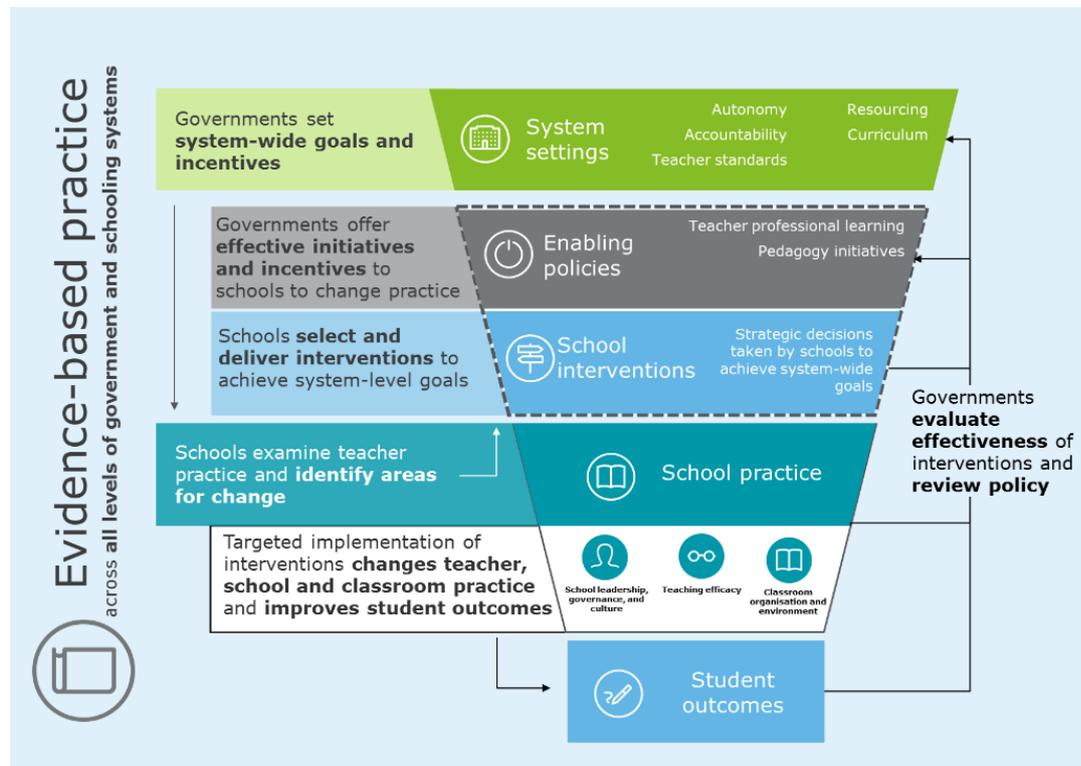
However, schools do not determine and implement changes to these drivers in isolation. Schools develop an understanding of the appropriate choices to make in response to identified challenges from a number of sources. Teachers are influenced by the practices that they have been taught in their initial training, and that they have developed over the course of their career. They may be influenced by their own research into academic evidence. Schools also have the option of procuring professional learning from private and public providers, and participating in system-led initiatives.

Figure 5.1 below presents an emerging framework for understanding the factors which influence school practice, and the role of government in setting system settings and enabling initiatives which provide the necessary conditions for schools to identify and invest in high quality practice.

- Governments establish the broad architecture of the schooling system, determining:
 - The resourcing provided to schools;
 - The content taught in schools (curriculum);
 - The parameters through which schools make decisions (autonomy) and report on practice and outcomes (accountability); and
 - The requirements individuals need to meet to become, and continue to work as, teachers.
- Governments also directly offer options to schools for improving teacher practice, such as changes to pedagogical approach for particular subject areas.
- Schools, operating within these parameters, make decisions about where they should focus their attention in a process of continual improvement. They do this by selecting interventions from government as well as from private providers (companies offering curriculum and practice material for purchase by schools), and from the practices of other teachers and other schools.
- Targeted implementation of these interventions leads to improvements on the key drivers of school quality, which in turn leads to outcomes improvement.

- At all levels (from the system to the school), evidence underlies the selection of appropriate interventions, and those interventions are evaluated for their effectiveness, helping to build the evidence base into the future.

Figure 5.1: Emerging framework for the role of government in the schooling system



Source: Deloitte Access Economics

Most importantly, in this framework, governments will often not seek to prescribe changes to school practice directly. The system instead has an important role to play in holding schools accountable for pursuing and achieving improvements in outcomes and other goals.

In this context, system settings are a necessary pre-condition for establishing an environment where schools identify areas for change, and invest in initiatives that meets the needs of students in their local context. Importantly, policymakers should **seek to demonstrate a link** between any proposed initiative, and drivers that are shown to most significantly contribute to variation in outcomes.

For many initiatives, this link is intuitive: for instance, teacher standards and initial teacher education accreditation helps to set a standard of professionalism and practice, and provide incentives for continual development across the teaching profession. A more indirect link can be drawn between the key drivers and the curriculum which educators use to guide their teaching practice. System-wide focuses on accountability and decentralisation of schooling policy helps to create a schooling system where teachers can pursue effective practices, though they do not direct those practices themselves.

Governments also play a key role in the improvement of school quality through:

- Using funding and grants to provide particular programs and incentivise certain practices in schools;
- Demonstrating the link to improvements in teaching practice before investing in initiatives intended to drive improvements in outcomes;
- Setting a strategic and long-term focus on the outcomes impact of sustained changes to practice at the classroom level;

- Relying on, and encouraging the use of, evidence-based interventions across Australian classrooms; and
- Collecting and sharing data and evidence on how interventions result in improved practice.

Funding and incentives

Although state and territory governments ultimately have responsibility for government schools in Australia's federal system, the Australian Government provides a substantial part of funding for schools: on average, close to a fifth of all public funding for schools. The Australian Government has also announced its intention to fund 20% of the Schooling Resource Standard for all government schools by 2027.

The Australian Government also offers funding incentives to state and territory governments implementing particular evidence-based reforms, through its Quality Schools, Quality Reforms initiative and through National Partnership Agreements (such as those supporting the implementation of autonomy measures through the Commonwealth Independent Public Schools initiative). These levers allow the Australian Government to encourage evidence-based reforms by state governments and at a school level.

Demonstrating a link between initiatives and improvements in practice

The framework and evidence developed through this study may be used to inform evaluation frameworks to be used by Australian jurisdictions to demonstrate a link between new policies and the drivers of school quality. This may include a framework which requires policy-makers to provide evidence of a link between policies—including those relating to resourcing—and effective school and classroom practice, measured at the level of the classroom.

Strategic and long-term focus

Alongside evaluations of individual programs, governments must make improvements in school quality a strategic priority. In many cases, initiatives are not evaluated against the existing data sets that these tests provide. Some initiatives will take more than a semester to evaluate—indeed, some may take many years to register a measurable impact (for instance, changes to initial teacher education). As a result, any implementation of initiatives—and evaluation of those initiatives—must be *sustained* to permit a true picture of effectiveness for school quality measures to emerge.

Evidence-based interventions

Government can also play a central role in *curating and evaluating the evidence base* which schools draw upon when making decisions about their practice and management.

This framework emphasises the role of government in monitoring and evaluating the effectiveness of schooling systems, enabling initiatives and practice to inform the system settings and enabling initiatives which guide school practice, while simultaneously holding the system accountable for driving improvements in student outcomes. An emphasis on evidence-based practice underpins all aspects of decision making within the system.

Current examples of best practice in collating such an evidence base include the Education Endowment Foundation's *Teaching and Learning Toolkits* (see Appendix B). Governments may then offer implementations of these initiatives to the schools that need them, or offer incentives which encourage teachers to acquire knowledge in line with the current state of thinking (by enabling teacher professional learning).

Collecting and sharing data evidence

The Productivity Commission (2016) observed that measuring outcomes in the education system can be done well with large scale datasets and simple data analysis. NAPLAN, PISA and TIMSS are already providing useful data to schools and policy makers on student outcomes. The most significant gap in Australia's education evidence base, according to the Commission, is on the impact or effect of particular initiatives. These questions require "a bottom-up approach, using small scale research projects and datasets that are often question-specific and apply sophisticated quantitative research methods." (Productivity Commission, 2016). Governments have a key role to play in developing this evidence base of *practice* to supplement the existing datasets relating to *outcomes*.

Notably, while datasets like TIMSS and PISA provide critical evidence on the contribution of teacher and school practice and management to student outcomes, they do not identify student or schools and so cannot be used to monitor the outcomes of particular schools or students—and therefore cannot be linked to specific policies for the purpose of evaluation and monitoring. The development of a similar dataset, that captures a level of detail in practice and performance, but may also be used for the purposes of monitoring and evaluation, would go some way in filling this gap. Indeed, this study has demonstrated the analytical value of such a dataset in terms of the ability for policy makers to examine and measure the key drivers of student outcomes and school quality across the Australian schooling system.

5.1.2 The current state of schooling policy in Australia

Commonwealth, state and territory governments establish the broad architecture for the operation of a largely decentralised schooling system. In Australia, the largest part of responsibility for schools is held by state and territory governments. In theory, differences between the policies in each jurisdiction could be linked to differences in outcomes. This could provide an indication that particular system-wide approaches to schooling are contributing to student achievement.

To unpack the relationship between system-level differences and variation between jurisdictions, a desktop review of 'grey literature' has identified the most significant schooling system changes in Australia over the past 15-20 years.²⁹ The detail of this review is contained in Appendix D. The key areas of focus identified by this review were:

- The implementation of the Australian Curriculum in each jurisdiction;
- Differences in the degree of autonomy offered to schools in financial and facilities management, determining their staffing profile, and filling vacancies with staff;
- The implementation of standardised testing initiatives prior to the National Assessment Plan for Literacy and Numeracy (NAPLAN);
- Historical differences in industrial arrangements (particularly the length of time over which teacher salaries reach their peak, and arrangements for 'skilled teaching positions');
- Teacher education course accreditation and teacher standards;
- Teacher professional learning; and
- Pedagogy initiatives.

The results from the research in Appendix D **do not reveal any consistent relationships** between particular aspects of system policy and overall performance, when analysed at a jurisdictional level. In some respects, this is not surprising, as it reaffirms the observation that it is practice at the school level, rather than distinct system level policies, which have the most direct relationship with student outcomes. It also demonstrates that there is no apparent 'silver-bullet' policy prescription—at the system level—which can be associated with higher (or lower) levels of performance. Rather, it is a range of enabling policies and initiatives which combine to provide the right environment to allow for effective classroom environment and teaching practice to eventuate.

This is not to suggest that governments—as makers of system level policy—cannot materially influence school quality and student outcomes in Australia. Indeed, many of these enabling policies and initiatives have clear conceptual links to the drivers of school quality which have been found to be most significant in driving student outcomes.

Commonwealth, state and territory governments are highly active in the schooling space. A preliminary review by the Department of Education for this project found more than 300 discrete policy initiatives, across a number of areas, over the past 10-15 years or more.

It is difficult to examine the impact of these varied initiatives at a system-wide level. This does not mean that they are not worthwhile investments for Australia's complex and diverse schooling system to be making. From a measurement perspective, there are a number of reasons why this analysis cannot observe the impact of such initiatives on student outcomes:

²⁹ This research included Commonwealth, state and territory education department policy and procedure documents, policy reviews from the peer-reviewed literature and bodies such as the Australian Centre for Educational Research (ACER) and policy instruments such as industrial agreements.

- They may be aimed at **intermediate outcomes** such as student access, participation and engagement. Although these are necessary preconditions to student success, it would not be expected that such interventions to have an immediate system-wide impact.
- They may be aimed at **broader school system goals** than literacy, numeracy and science. Civics, languages other than English, information technology and arts education all form part of a school-level education in developed nations; however, initiatives in these categories are unlikely to have a direct effect on literacy and numeracy performance.
- They may be targeted at **individuals, or particular groups of students**. Scholarship programs, for instance, undoubtedly have a personal impact on recipients, but their effects are unlikely to be represented in system-wide results. Rather, their aim is to enable individual students' achievements or particular equity related goals.
- Most programs are **not implemented across all schools at once**. Indeed, very few initiatives are likely to lead to across-the-board improvement in every school the same way. Government-led initiatives, sensibly, are piloted with small groups of schools, and used in schools with recognised areas of local need.
- Initiatives may also have considerable **impact lags**. In particular, those relating to changing the process of initial entry into the teacher profession – be it changes to initial teacher education, or incentives in industrial arrangements – rely on a long-term change to the teaching workforce, meaning their effect will not be clear from a system-level analysis.
- The exercise of selecting policies that may, or may not, have had an impact on PISA or TIMSS results requires **subjective decisions** about the priority of particular settings within schooling systems, and a level of knowledge about the implementation of initiatives and student outcomes that is not presently available at a school level.

5.2 The importance of evaluation

Ultimately, the sheer number of initiatives being taken by state and territory governments limits the ability of empirical analysis to isolate their impact without direct observations of school and classroom level data. Evaluating the effectiveness of particular school-level interventions, from a system perspective, requires at a minimum an understanding of:

- The types of intervention implemented;
- The school and group of students subject to the intervention; and
- The outcomes of students before and after the intervention.

This is particularly relevant knowing that teaching practice is the most significant driver of school quality. Although government policy may steer schools in a particular direction, the evidence from this study's analysis of PISA and TIMSS suggests that through the practice of individual teachers that particular policy initiatives lead to changes in outcomes. Governments play an important role in setting a strategic direction for schools and creating the culture in which teachers operate, but ultimately the methods used by teachers in classrooms are determined by those teachers.

Nonetheless, the limitations preventing this type of evaluation demonstrate that further evidence and analysis is required to understand the link between:

- The drivers of student outcomes and school quality that are understood to matter most;
- The levers which policy-makers can use to influence these aspects of schooling; and
- The overall performance of education systems, in terms of student outcomes.

Australian states and territories have been collecting broad-based outcomes data on numeracy and literacy for many years now. This is now done in a standardised manner nationally, through the NAPLAN process. This is intended to provide an incentive to monitor individual students' progress, benchmark school performance and create competition between schools. However, as the Productivity Commission (2016) noted, this type of 'top-down' evidence alone does not ensure that Australian schools can realise gains in outcomes. The evidence base in the Australian education community, according to the Commission, lacks a 'bottom-up' assessment of policies, programs and teaching practices.

Detailed individual evaluations of particular policies exist and can demonstrate whether particular 'levers' have worked in improving student outcomes in particular contexts. However, these provide

limited strategic insight into how improvements may be made within the context of a consistent, strategic, overarching policy framework, which is oriented towards the drivers of school quality.

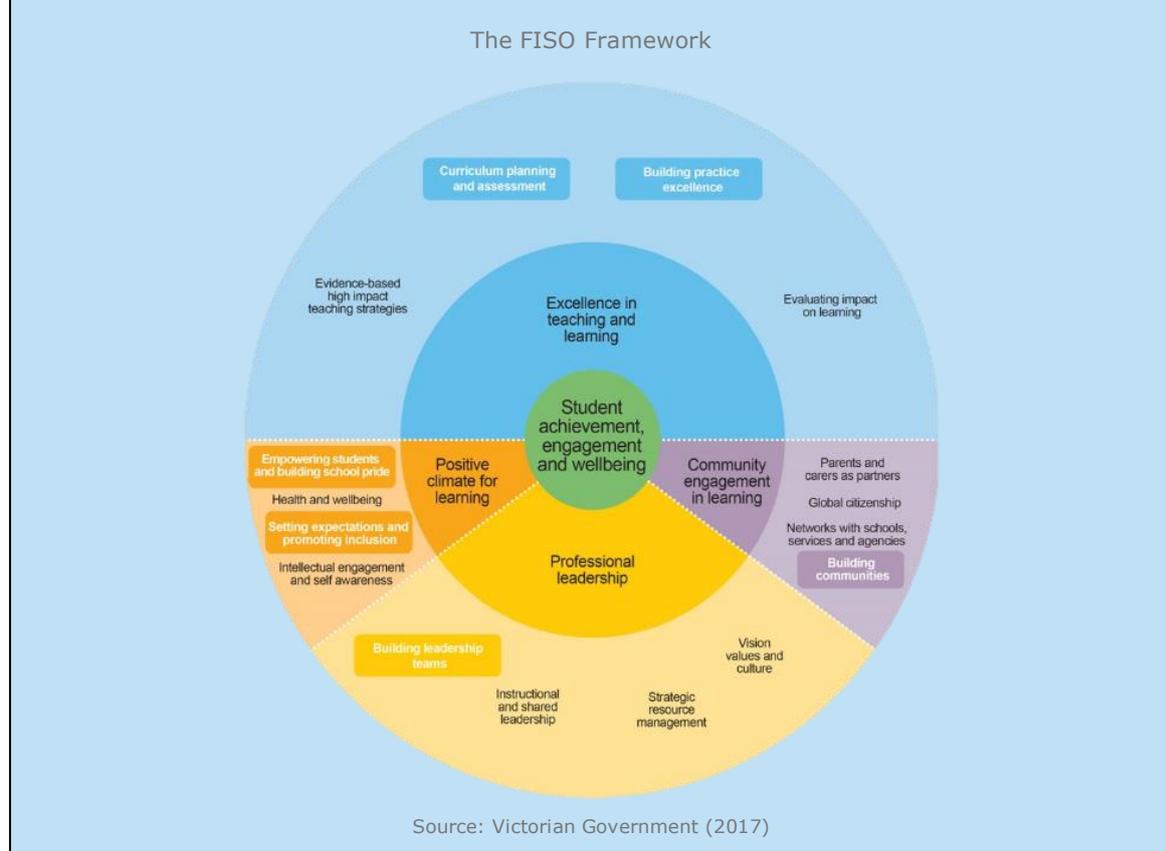
At the centre of such strategic policy direction is a connection between the aspects of school practice that are known to matter and the role of government in determining system level policy. While more research and analysis will be required to better understand this link, which falls outside of the scope of this study, the measurement framework and initial analysis developed here provides a foundation for further research to be completed.

Case study – Framework for Improving Student Outcomes (FISO)

In Victoria, the Framework for Improving Student Outcomes (FISO) provides information, resources and support for schools to identify and implement strategies that enable school improvement. In line with the indications of our empirical analysis, it recognises that schools are best placed to understand what works for their school, and aims to balance school-level decision making with a need for evidence, research and best practice backing to any initiatives.

The FISO is comprised of four headline priorities: excellence in teaching and learning, professional leadership, a positive climate for learning and community engagement in learning which are based on robust evidence of the drivers of student outcome, in proportion to their relative importance. Among these headline priorities, there are 16 initiatives, most of which focus on the excellence in teaching and learning priority.

This case study provides an example of how evidence of the drivers of student outcomes in Australia, within a consistent measurement framework, can be used to guide school’s strategic decision making towards the improvement of student outcomes.



5.2.1 Possible implications for policy-makers

The research outlined in Appendix D has identified a significant and diverse range of schooling interventions which have occurred in Australian schools over the past 10-15 years. The findings from this study may be used to support Australian jurisdictions in demonstrating a link between

new policies and known drivers of school quality. These findings may also provide the basis for developing a new evidence accountability framework which requires policy-makers to provide evidence of a link between policies—including those relating to resourcing—and effective school and classroom practice, measured at the level of the classroom.

This study demonstrates how a system-wide evidence base of practice, policy and performance in Australia may be used to provide a robust evidence base to inform policy. Building on this work, next steps for government may include:

- Developing a consistent and system-wide evidence base for evaluating the impact of government policies and programs on student-level outcomes;
- Maintain key long term strategic goals which are informed by a robust evidence-base;
- Establishing accountable and transparent schooling systems; and
- Continuously evaluating and refining Australian policy and practice.

Developing a consistent and system-wide evidence base

While diversity and complexity in policy design and application need not be a shortcoming, the lack of a consistent and universal basis for evaluating the impact of government policies and programs on student-level outcomes means there exists little capacity to assure Australia is on a path towards overall school improvement.

For this task, there are abundant evidence bases about both outcomes and practice from professional bodies, private educational businesses, academics and government. However, it is apparent that a given intervention that has been demonstrated to improve student outcomes in one context will not necessarily produce the same result in other contexts.

As a result, all parts of the schooling system (from the National and State governments to schools themselves) must play a role in continual evaluation of school practice and building the broader evidence base. This is in line with the recommendations of the Productivity Commission (2016) in its recent Inquiry Report into the National Education Evidence Base. The suggested approach of the Commission combines 'top-down' transparent assessment and collection of outcomes data with 'bottom-up' evaluation of programs. Combined, the availability of data on both student outcomes and practice will help to facilitate conversations between levels of government and schools on how to drive student outcomes.

Maintaining long-term strategic goals

In particular, the maintenance of a clear strategic policy direction is critical in ensuring that the development of new policies and initiatives are consistently and clearly oriented towards the drivers of practice which matter for schools. **This report reinforces the most crucial areas of school practice that have the most potential to influence student outcomes.** A strategic policy direction built on this evidence would be expected to direct the focus of all areas of policy to both *proven effective initiatives*, and to *continual evaluation* of policies, that relate to these drivers.

Establishing accountable and transparent schooling systems

In order to evaluate the impact of initiatives and policies enacted across Australia's schools, governments must take responsibility for establishing accountable and transparent schooling systems. Collecting outcomes data in a granular but broad manner, and sharing it nationally, will help to facilitate analysis and discussion about the impact of initiatives.

Continuous evaluation and refinement

Beyond just collecting data on outcomes, it is crucial for data to be used at all levels of the schooling system to critically reflect on, and refine, the practices used in schools. The OECD, in its Review on Evaluation and Assessment Frameworks for Improving School Outcomes, observed that evaluation is needed at student, teacher, school and system levels, but noted universal challenges to the implementation of effective evaluation systems in schools. In particular, the goals of each form of evaluation need to be aligned, to ensure teachers and schools are pursuing system-level goals in a strategic and focused manner. In this area, Australian schools still have some progress to be made – for example, a review of teacher effectiveness evaluation frameworks found that

there was still room to develop a nationally consistent approach (University of Melbourne Graduate School of Education, 2017).

Below, a case study of schooling system reform in Ontario, Canada, which implemented a system-wide approach to improving student outcomes is detailed.

Case study of system reform – Canada, Ontario

In the early 2000s the school education system in Ontario, Canada was shown to perform poorly on a range of national indicators. A new provincial government elected in 2003 prioritised education transformation and change.

The reform framework was characterised by a deliberately contained and limited suite of key objectives and a coherent structure of concerted support—focus, build relationships, persist, develop capacity, and spread quality implementation. The system explicitly committed to raising performance outcomes for all students and closing achievement gaps between all groups.

Two major initiatives were accordingly pursued by the Ontario Ministry of Education over the reform time period:

1. The Literacy and Numeracy initiative aimed to increase reading and mathematics outcomes in elementary schools. Described as a capacity-building strategy, the initiative succeeded in raising the average pass rate in provincial exams from approximately 55% (2003) to approximately 70% (2010) in reading, mathematics and writing in grade 3. Similar gains of about 10-12 percentage points were apparent in the same subjects in grade six.

2. The Student Success initiative aimed to increase the high school graduation rate to 85%. By identifying students at risk early, funding a “student success officer” in each school, and creating programmes of “credit recovery” through which students could make up the parts of courses that they failed, the graduation rate increased from 68% to 79% over several years.

Key success factors identified for enabling the Ontario system to achieve progress on key indicators included the consistent application of centrally-driven pressure for higher results, combined with extensive capacity building and a climate of relative trust and mutual respect.

Strategic measures associated with the success of this initiative included:

- **Strategies directly focused on improving teaching practice.**
- Careful and detailed attention to implementation, along with opportunities for teachers to practice new ideas and learn from their colleagues.
- A single integrated strategy and one set of expectations for both teachers and students.
- Support from teachers for the reforms.

The continued use of this new methodology, coupled with expanded evidence of practice and policy across Australian schools, enhances the depth and scope of evidence and insights available to the Department, and provide the basis for developing a framework for developing consistent national evidence that links policy, practice and performance—allowing for greater transparency and accountability in the policy interventions that are taking place across schools and jurisdictions.

The findings from this study may be used to support Australian jurisdictions in demonstrating a link between new policies and known drivers of school quality. These findings also provide the basis for developing a new evidence accountability framework which requires policy-makers to provide evidence of a link between policies—including those relating to resourcing—and effective school and classroom practice, measured at the level of the classroom.

5.3 Conclusions and future research directions

The empirical methodology developed through this study, and the underpinning measurement framework, is intended to provide the Department with a detailed and impactful evidence base to inform future directions for government. In particular, it has answered the key research questions established for this study by providing:

- A framework and methodological approach which provides greater **clarity** in understanding the drivers of student outcomes and school quality in Australia;
- A robust, detailed and *Australian specific* evidence-base which builds **confidence** for policy-makers in understanding the key factors which contribute to student outcomes in Australia
- An emerging framework to support the **strategic focus** and direction for policymakers when considering the role of government in driving improvement in Australia's school education system.

Future research may build on and refine this methodology towards providing further evidence and insight. This future research may expand on this study by, among other things:

- Expanding the scope of the empirical analysis to examine the effects of different drivers of school quality over time (for example, by mapping key PISA and TIMSS questions across years). This will be an important test of the external validity of the findings provided in this initial application of the established methodology.
- Incorporating (more fully) additional student outcome dependent variables, incorporating aspects of student engagement and wellbeing.
- Adding further causal structure to the empirical analysis to understand how different drivers of quality affect each other, and then subsequently drive student outcomes (for example, by estimating the link between school leadership and teaching practice).
 - A possible causal structure may incorporate the following nested components (which affect each other in turn, with the final component driving student outcomes):
 1. System policy settings (to the extent they can be identified in PISA and TIMSS)
 2. School leadership and management
 3. Teacher attributes
 4. Classroom organisation
 5. Classroom environment
 6. Teaching practice
 7. Student engagement and wellbeing
 8. Student academic achievement
 - This more complex structural analysis may provide further insights into how different drivers of school quality influence each other, and provide a more complete link between the role of system level policies, school practice, teacher effectiveness, and student outcomes.
- Further exploration of the impact of prior performance on outcomes. The use of other data sources such as NAPLAN, or more detailed analysis of measures of student self-efficacy, may be incorporated into the modelling to examine how significant the omission of prior performance is on the findings of this study.
- The TIMSS dataset may be used to explore the importance of variation within schools (and across classrooms) as well as across schools, to better understand the relative effects of teaching practice and school quality, and how teaching practice contributes to overall measures of school quality.
- More detailed comparisons can be made across different learning domains (such as reading, maths and science) to understand how different aspects of school practice and management can affect influence performance in different areas. (For example, do teachers 'matter' more for maths than science?)
- Considering the heterogeneity of estimated effects across the distribution of student outcomes, and for students with different characteristics. For example further analysis may explore whether the effects of aspects of teaching practice matter more or less for low and/or high socio-economic status schools, or in different schooling systems/jurisdictions. From a practical standpoint, this may involve enhancements to the regression models that include interaction terms, random forests, generalised additive models, etc.
- International data may be incorporated into the methodology to expand the analysis of Australia's performance relative to other countries, and provide insights into how the identified drivers of school quality vary across countries, in their nature and relative effects.

References

- AAP – Australian Associated Press (2016) *WA's 2016 NAPLAN results best in Australia*.
<http://www.perthnow.com.au/news/western-australia/was-2016-naplan-results-best-in-australia/news-story/74884a860d7831137e8b30816e673aa0>
- ACARA - Australian Curriculum and Reporting Authority (2014) *State and Territory Implementation of the Foundation to Year 10 Australian Curriculum*.
https://acaraweb.blob.core.windows.net/resources/State_and_Territory_F-10_Australian_Curriculum_Implementation_Timelines_August_2014.pdf
- ACT Directorate of Education and Training (2004) *Report on Evaluation of School Based Management*.
http://www.education.act.gov.au/_data/assets/pdf_file/0009/17982/sbmevaluationreport.pdf
- ACT Directorate of Education and Training (2012) *Procedures for filling classroom teacher vacancies*.
http://www.det.act.gov.au/_data/assets/pdf_file/0015/290310/120322_Filling_Classroom_Teacher_Vacancies_-_AEU_Exec_endorsed.pdf
- Alexander, Karl L., Doris R. Entwisle, Dale A. Blyth, and Harriette Pipes McAdoo. "Achievement in the First 2 Years of School: Patterns and Processes." *Monographs of the Society for Research in Child Development* 53, no. 2 (1988): i. doi: 10.2307/1166081.
- Anderson, M.; Gronn, P.; Ingvarson, L.; Jackson, A.; Kleinhenz, E.; McKenzie, P.; Mulford, B.; Thornton, N. (2007) *OECD Improving School Leadership Activity – Australia: Country Background Report*. <https://www.oecd.org/edu/school/39967643.pdf>
- ARACY – Australian Research Alliance for Children & Youth (2012) *Parental engagement in learning and schooling: lessons from research*. https://www.aracy.org.au/publications-resources/command/download_file/id/7/filename/Parental_engagement_in_learning_and_schooling_Lessons_from_research_BUREAU_ARACY_August_2012.pdf
- Australian Bureau of Statistics (2015) *Frameworks for Australian Social Statistics*.
<http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4160.0.55.001~Jun%202015~Main%20Features~What%20is%20a%20framework%3f~6>
- Australian Education Union (2017) *In-principle Victorian Government Schools Agreement 2017*.
<https://www.aeuvic.asn.au/sites/default/files/In-principle%20Victorian%20Government%20Schools%20Agreement%202017.pdf>
- NSW Department of Education (2015) *Standards Based Pay for Permanent and Temporary Salaries*. <https://www.det.nsw.edu.au/media/downloads/about-us/news-at-det/announcements/russell-lea-consultation/19602AwardsMaterialSalaryScaleNovember2015.pdf>
- Australian Government (2014a) *Project Agreement for Independent Public Schools Initiative – Northern Territory*.
http://www.federalfinancialrelations.gov.au/content/npa/education/project-agreement/ind_public_schools_NT_ips.pdf
- Australian Government (2014b) *Project Agreement for Community Empowered Schools – Tasmania*. http://www.federalfinancialrelations.gov.au/content/npa/education/project-agreement/ind_public_schools_Tas_ips.pdf

- Australian Government (2017) *Project Agreements – Education*.
<http://www.federalfinancialrelations.gov.au/content/npa/education.aspx>
- Australian Institute of Family Studies (2015) *Working with culturally and linguistically diverse (CALD) adolescents*. <https://aifs.gov.au/cfca/publications/working-culturally-and-linguistically-diverse-cald-adolescents/reports-policy-papers>
- Bandura, Albert (1997), *Self-efficacy: The Exercise of Control*. New York: Freeman
- Belfield, Clive R., and Henry M. Levin. "The Effects of Competition Between Schools on Educational Outcomes: A Review for the United States." *Review of Educational Research* 72, no. 2 (June 1, 2002): 279–341. doi:10.3102/00346543072002279.
- Betts, Julian R., Andrew Zau, and Lorien Rice. *Determinants of student achievement: New evidence from San Diego*. San Francisco: Public Policy Institute of California, 2003.
- Biddle, N. Cameron, T. (2010). Potential factors influencing Indigenous education participation and achievement. Centre for Aboriginal Economic Policy Research, Australian National University)
- Brookover, W.B., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference*. New York: Praeger.
- Burgess, S (2016); *Human Capital and Education: The State of the Art in the Economics of Education*, IZA Discussion Paper, No. 9885, April 2016.
- Burke, Mary A., and Tim R. Sass. "Classroom Peer Effects and Student Achievement." *Journal of Labor Economics* 31, no. 1 (2013): 51 – 82.
- Caldwell, B. (2015) *School Autonomy and Student Achievement – case studies in Australia*.
<http://educationaltransformations.com.au/wp-content/uploads/School-Autonomy-and-Student-Achievement-Australian-Cases.pdf>
- Centre for Education and Statistics (CESE). (2017). *The Teaching and Learning Toolkit*. New South Wales Government
- Centre for Education and Statistics (CESE), Lu, L. and Rickard, K. (2014). *Value added models for NSW government schools*. Centre for Education and Statistics (CESE). New South Wales Government
- Coe, R., Aloisi, C., Higgins, S., Elliot Major, Lee. (2014). *What makes great teaching? Review of the underpinning research*. Centre for Evaluation & Mentoring
- Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. Teachers College Press.
- Day, Christopher., Sammons, Pam., Hopkins, David., Harris, Alma., Leithwood, Ken., Gu, Qing., Brown, Eleanor., Ahtaridou, Elpida and Kington, Alison., (2009) *The Impact Of School Leadership On Pupil Outcomes*, University of Nottingham Research Report DCSF-RR108.
- Deloitte Access Economics, 2014, *Effective investment to improve educational outcomes*, Department of Education and Early Childhood Development
- Department of Education and Early Childhood Development (DEECD) Victoria. (2007). *Value-added measures for school improvement*, no.13, Victorian Government, Melbourne.
- Dolton, P., Chevalier, A. & McIntosh, S. (2001). *Recruiting and retaining teachers in the UK: An analysis of graduate occupation choice from the 1960s to the 1990s*. London: Department of Education and Science.

- Dornyei, Z. (2001). *Teaching and researching motivation*. New York: Longman
- Duncan, G.J., Dowsett, C.J., Claessens, A., Mangnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446
- Edmonds, R.R. (1979a). Effective schools for the urban poor. *Educational Leadership*, 37 (1), 15-17.
- Edmonds, R.R. (1979b). Some schools work and more can. *Social Policy*, 9, 28-32.
- Field, S, Kuczera, M & Pont, B 2007, *No more failures: Ten steps to equity in education*, OECD, Paris
- Forgasz, Helen J.; Leder, Gilah C.; Taylor, Calvin. (2007), *Research versus the media: Mixed or single-gender settings?* Monash University – Presentation Paper for AARE Annual Conference, Fremantle
- Fuchs, T., Woessmann, L. (2004). What accounts for international differences in student performance? A re-examination using PISA data. Leibniz Information Centre for Economics
- Gonski, D. M. (2011). *Review of funding for schooling: final report*. Australia. Department of Education, Employment and Workplace Relations. Review of Funding for Schooling.
- Grant, G. (1988). *The world we created at Hamilton High*. Cambridge, MA: Harvard University Press
- Griffin, P., Woods, K., & Nguyen, T. (2005). *An environmental scan of tools and strategies that measure progress in school reform*. Report to the Department of Education and Training, Melbourne, Victoria.
- Hanushek, E. A. (2003). The Failure of Input-based Schooling Policies*. *The economic journal*, 113(485), F64-F98.
- Hanushek, Eric A. "The Economic Value of Higher Teacher Quality." *Economics of Education Review* 30, no. 3 (June 2011): 466–79. doi:10.1016/j.econedurev.2010.12.006.
- Hanushek, Eric A., Susanne Link, and Ludger Woessmann, 2013 "Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA." *Journal of Development Economics* 104 (September 2013): 212–32. doi:10.1016/j.jdeveco.2012.08.002.
- Hattie, J. (2009). *Visible Learning, A synthesis of Over 800 meta-analyses relating to achievement*, Routledge, Oxon
- Hay, I, Castleton, G, Callingham, R, Edmunds, B, Fielding-Barnsley, R, Moss, T & Grimbeek, P (2011) *Advancing Literacy in Tasmanian Primary Schools*, University of Tasmania & Tasmanian Department of Education, Hobart, Tasmania.
- Heckman, JJ and Kautz, T (2013) *Fostering and measuring skills: Interventions that improve character and cognition* (No. w19656), National Bureau of Economic Research.
- Hemmings, Brian, Peter Grootenboer and Russell Kay (2010) *Predicting mathematics achievement: the influence of prior achievement and attitudes*. *International Journal of Science and Mathematics Education* 9(3) (June 2011), 691-705.
- Heyneman, Stephen P., and William A. Loxley. "The effect of primary-school quality on academic achievement across twenty-nine high-and low-income countries." *American Journal of sociology* (1983): 1162-1194.

- Hill, P. W. (1995). Value-added measures of achievement. Incorporated Association of Registered Teachers of Victoria (IARTV) Seminar Series, no. 44.
- Ingvarson, L., Rowe, K. (2007). Conceptualising and evaluating teacher quality: Substantive and methodological issues. Australian Council for Educational Research
- Ingvarson, L; Kleinhenz, E; Wilkinson, J (2007) *Research on performance pay for teachers*.
<http://research.acer.edu.au/workforce/1/>
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gentis, H., Heynes, B., & Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Jensen, B. (2010). Investing in our teachers, investing in our economy. Grattan Institute
- Jensen, B. (2014). Making time for great teaching. Grattan Institute
- Kang, Changhui. "Classroom Peer Effects and Academic Achievement: Quasi-Randomization Evidence from South Korea." *Journal of Urban Economics* 61, no. 3 (May 2007): 458–95. doi:10.1016/j.jue.2006.07.006.
- Kilvert, P. (2001) *Partnerships 21*.
<https://ehlt.flinders.edu.au/education/iej/articles/v2n1/kilvert/kilvert.pdf>
- Krueger, Alan, and Diane Whitmore. The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR. NBER Working Paper. National Bureau of Economic Research, Inc, 2000.
<http://ideas.repec.org/p/nbr/nberwo/7656.html>.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implication for graduate student selection and performance. *Psychological Bulletin*, 127(1), 162-181
- La Paro, K. M., & Pianta, R. C. (2000) Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research*, 70(4), 443-484
- Lavy, V. and Sand, E. *On The Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Stereotypical Biases*, NBER Working Paper No. 20909.
- Lee, V. E., Dedrick, R. F., & Smith, J. B. (1991). The effect of the social organisation of schools on teacher efficacy and sense of satisfaction. *Sociology of Education*, 64(3), 190-208.
- Lightfoot, S. L. (1983). *The good high school: portraits of character and culture*. Basic Books, New York
- Linnenbrink, E. A. and Pintrich, P. R. (2003) The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading and Writing Quarterly*, 19: 119-137
- Liu, Y., Zumbo, B.D., & Wu, A.D. (2014). 'Relative importance of predictors in multilevel modelling', *Journal of Modern Applied Statistical Methods*, 13(1), 2-22.
- Maher, Marguerite., (2013), Early childhood project analysed within a model enhancing the self-efficacy of Indigenous people, *Education Papers and Journal Articles – University of Notre Dame Australia*
- Marks, H and Printy, S, 2003, Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*.

- McRae, D.; Ainsworth, G.; Groves, R.; Rowland, M.; Zbar, V (2001) *PD 2000 Australia: A National Mapping of School Teacher Professional Development*. <http://pandora.nla.gov.au/tep/21763>
- McRae, David., Ainsworth, Geoff., Cumming, Jim., Hughes, Paul., Mackay, Tony., Price, Kaye., Rowland, Mike., Warhurst, Joan., Woods, Davina., Zbar, Vic., (2000), What works? Explorations in improving outcomes for Indigenous students, Commonwealth Department of Education, Training and Youth Affairs (*Dissolved*)
- Meiers, M.; Ingvarson, L.; Beavis, A.; Hogan, J.; Kleinhenz, E. (2008) *An evaluation of the Getting it Right: Literacy and Numeracy Strategy in Western Australian schools*. http://research.acer.edu.au/policy_reform/2/
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of counselling psychology*, 38(1), 30-38
- Nakata, M., (2007), The Cultural Interface, *The Australian Journal of Indigenous Education* Volume 36
- NSW Board of Studies (2012) *Update on the Implementation of Australian Curriculum in NSW*. http://www.boardofstudies.nsw.edu.au/australian-curriculum/pdf_doc/120731-memo-to-principals-update-on-ac-implementation.pdf
- NSW Department of Education and Communities (2014) *Local Schools, Local Decisions – The Reform Agenda*. <https://www.det.nsw.edu.au/media/downloads/about-us/our-reforms/lslr/LSLD-overview-july2014.pdf>
- NSW Education Standards Authority (2017) *Minimum HSC Standard*. <https://educationstandards.nsw.edu.au/wps/portal/nesa/about/initiatives/stronger-hsc-standards/minimum-HSC-standard/multiple-opportunities>
- Odden, A. R., Picus, L. O., Archibald, S., Goetz, M., Mangan, M. T., & Aportela, A. (2007). Moving from good to great in Wisconsin: Funding schools adequately and doubling student performance. Consortium for Policy Research in Education, Wisconsin Center for Educational Research.
- OECD – Organisation for Economic Cooperation and Development (2016) *Education at a Glance 2016*.
- OECD (2013), *PISA 2012 Results: What Makes Schools Successful (Volume IV): Resources, Policies and Practices*, PISA, OECD Publishing.
- OECD (2013), *Key Findings from the Teaching and Learning International Survey (TALIS) – Australia*. <https://www.oecd.org/australia/TALIS-2013-country-note-Australia.pdf>
- Organisation for Economic Co-operation and Development (OECD). (2008). *Measuring improvements in learning outcomes: best practices to assess the value-added of schools*, Paris: OECD Publishing.
- Pajares, F (2003), *Self-efficacy beliefs, motivation and achievement in writing: A review of the literature*. *Reading and Writing Quarterly*, 19: 139:158.
- Parkville Global Advisory (2014), *National Evaluation for the Low SES National Partnership and the Literacy and Numeracy National Partnership – Impact Stage, Draft Final Report*.
- Productivity Commission (2016) *National Education Evidence Base*. <http://www.pc.gov.au/inquiries/completed/education-evidence/report/education-evidence-overview.pdf>.

- Queensland Department of Education and Training (n.d.) *Independent Public Schools – Policy Framework*. <http://education.qld.gov.au/schools/independent-public-schools/resources/ips-policy-framework.pdf>
- Raudenbush, S.W., & Willms, J.D. (Eds.). (1991). *Schools, classrooms and pupils: International studies of schooling from a multilevel perspective*. New York: Academic Press.
- Rowe, K. (2003). *The importance of teacher quality as a key determinant of students' experiences and outcomes of schooling*. Australian Council for Educational Research
- Rowe, K and National Inquiry into the Teaching of Literacy (Australia), "Teaching Reading" (2005). https://research.acer.edu.au/tll_misc/5
- Rutter, M., Maughan, B., Mortimer, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Somerset: Open Books.
- Samson, G. E., Graue, M. E., Weinstein, T., & Walberg, H. J. (1984). Academic and occupational performance: A quantitative synthesis. *American Educational Research Journal*, 21(2), 311-321
- Sanders, W. and J. Rivers (1996). "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement Summary of Findings. Knoxville, Tennessee."
- Sanders, W.L. (2000). Value-added assessment from student achievement data: opportunities and hurdles. *Journal of Personnel Evaluation in Education*, 14(4), 329-339.
- Sellstrom, E. (2006). Is there a "school effect" on pupil outcomes? A review of multilevel studies. *Journal of Epidemiology & Community Health*, 60(2), pp.149-155.
- South Australian Department for Education and Child Development (2015) *Policy – Recruitment and selection of teaching staff in schools*. https://www.decd.sa.gov.au/sites/g/files/net691/f/recruitment_teaching_schools_policy.pdf
- Stronge, J., Tucker, P. and Hindman, J. (2004). *Handbook for qualities of effective teachers*. Alexandria, Va.: Association for Supervision and Curriculum Development, Chapter 3.
- Teacher Education Ministerial Advisory Group (2014) *Action Now: Classroom Ready Teachers*. https://docs.education.gov.au/system/files/doc/other/action_now_classroom_ready_teachers_accessible.pdf
- The Education Endowment Foundation (EEF) 2014, Sutton Trust – EEF Teaching and Learning Toolkit. <http://educationendowmentfoundation.org.uk/toolkit/>
- University of Melbourne Graduate School of Education (2013) *Evaluation of the Independent Public Schools Initiative*. <https://www.education.wa.edu.au/documents/2548175/2664299/Evaluation+of+the+Independent+Public+Schools+initiative+%28Full+report%29.pdf/aafae012-c595-4447-943f-caf982c75785>
- University of Melbourne Graduate School of Education (2015) *Evaluation of the Implementation of the Australian Professional Standards for Teachers*. <https://www.aitsl.edu.au/docs/default-source/aitsl-research/report-2---school-and-organisational-implementation-of-the-standards---final.pdf?sfvrsn=2>
- University of Melbourne Graduate School of Education (2017) *Teacher effectiveness systems, frameworks and measures: a review*. <https://docs.education.gov.au/system/files/doc/other/teachereffectrpt2017.pdf>

Warwick, John. (2008), Mathematical self-efficacy and student engagement in the mathematics classroom, MSOR Connections Vol 8 No 3

Watson, Louise (2009) *Issues in reinventing school leadership: reviewing the OECD report on improving school leadership from an Australian perspective*. Leading & Managing, Vol. 15, No. 1, 1-13.

Western Australian Auditor-General (2011) *Right teacher, right place, right time: teacher placement in public schools*. https://audit.wa.gov.au/wp-content/uploads/2013/05/report2011_06.pdf

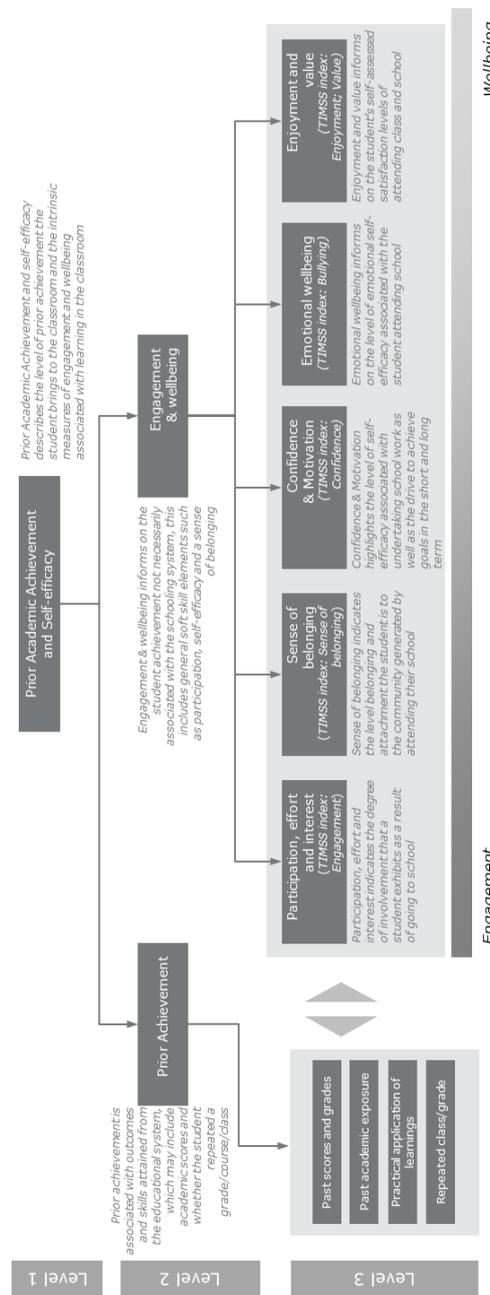
Western Australian Government (2013) *Major reforms to raise bar on literacy and numeracy*. <https://www.mediastatements.wa.gov.au/Pages/Barnett/2013/01/Major-reforms-to-raise-bar-on-literacy-and-numeracy.aspx>

Western Australian Government (2017) *OLNA results see high school bar raised in WA*. <https://www.mediastatements.wa.gov.au/Pages/Barnett/2017/01/OLNA-results-see-high-school-bar-raised-in-WA.aspx>

Willms (2004), Reading Achievement in Canada and the United States: Findings from the OECD Programme for International Student Assessment, Learning Policy Directorate Strategic Policy and Planning Human Resources and Skills Development, Canada

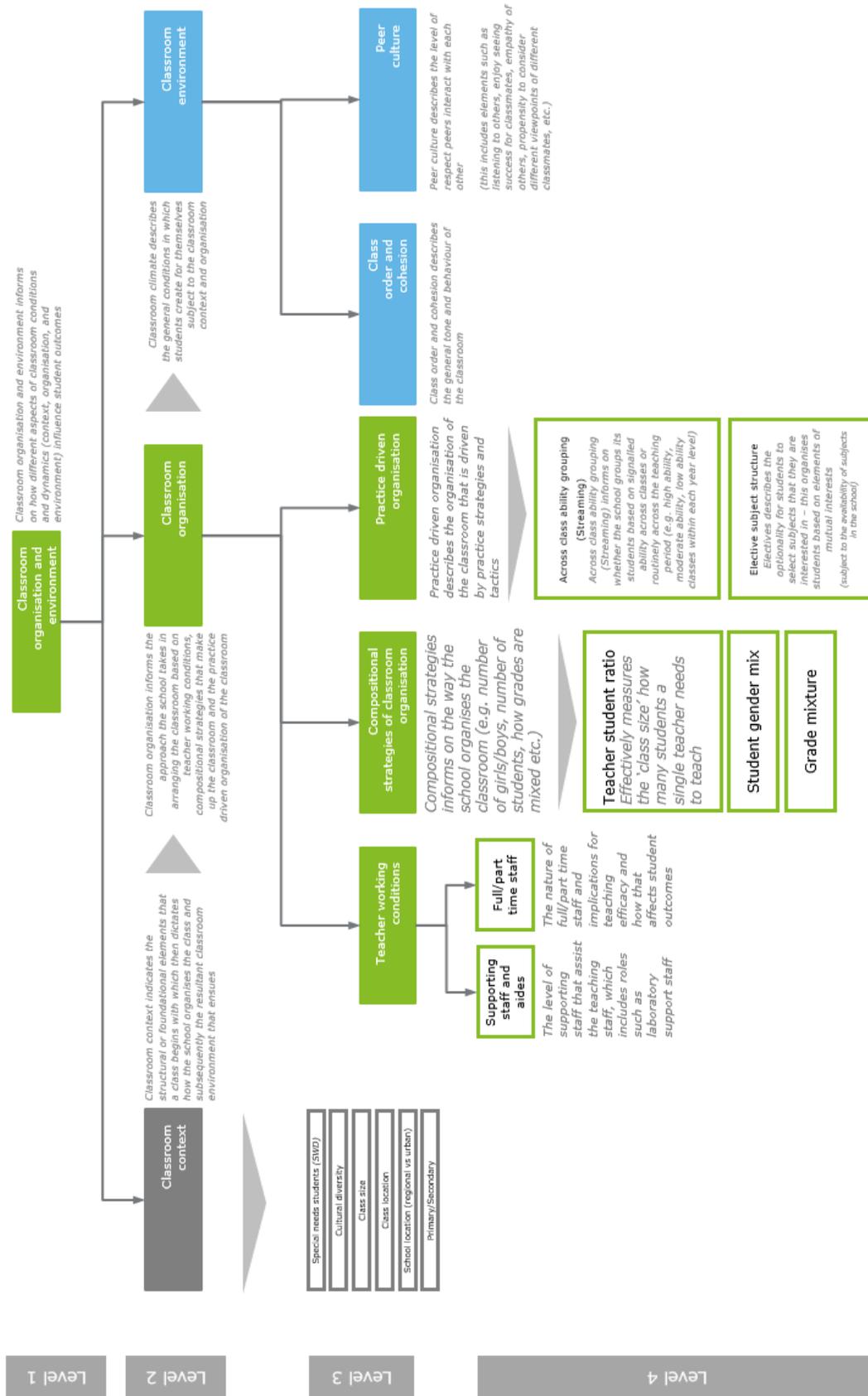
Appendix A: Detailed measurement frameworks, for mapping of PISA and TIMSS questions to themes

Chart A.1: Measurement framework for prior academic achievement and self-efficacy



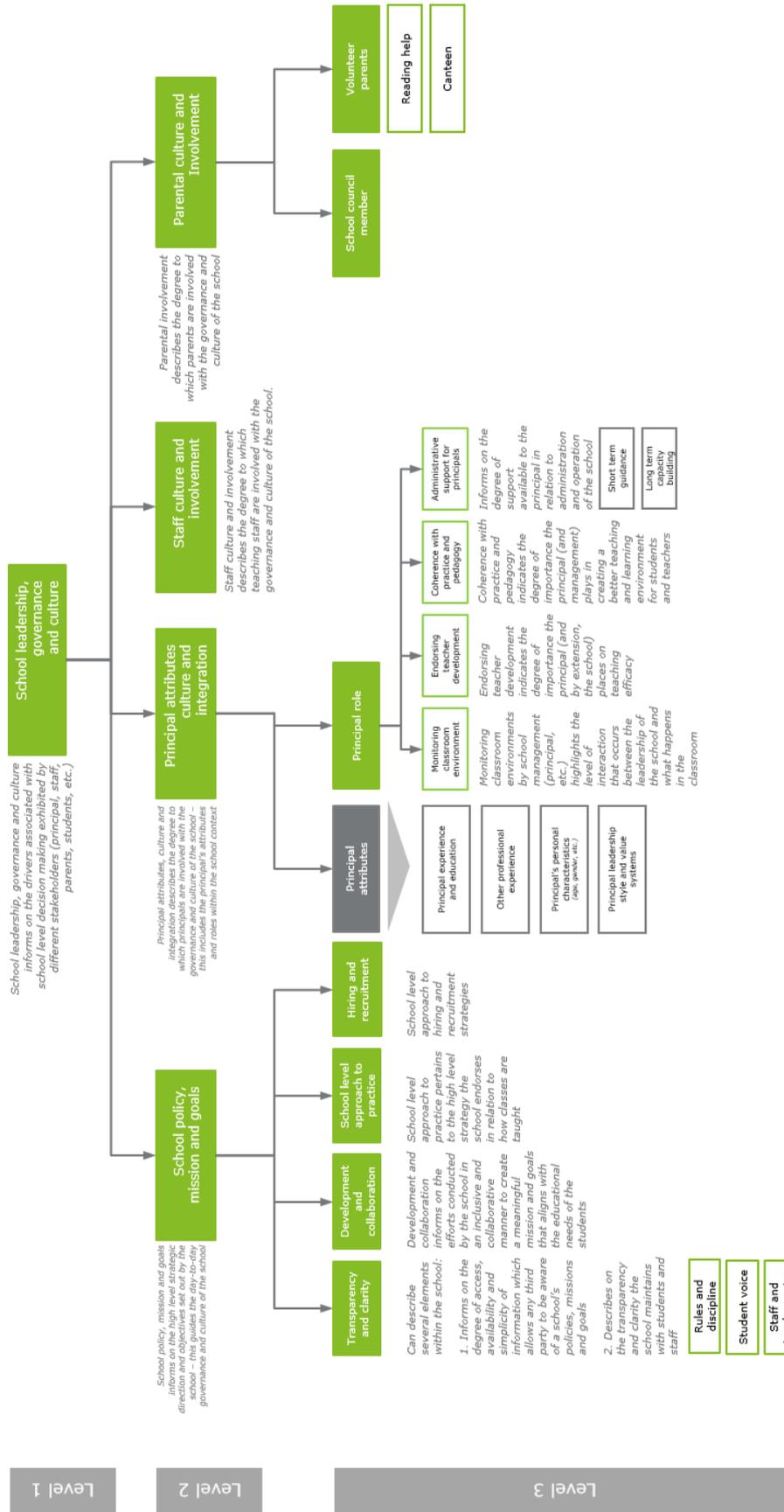
Source: Deloitte Access Economics

Chart A.3: Measurement framework for classroom organisation and environment



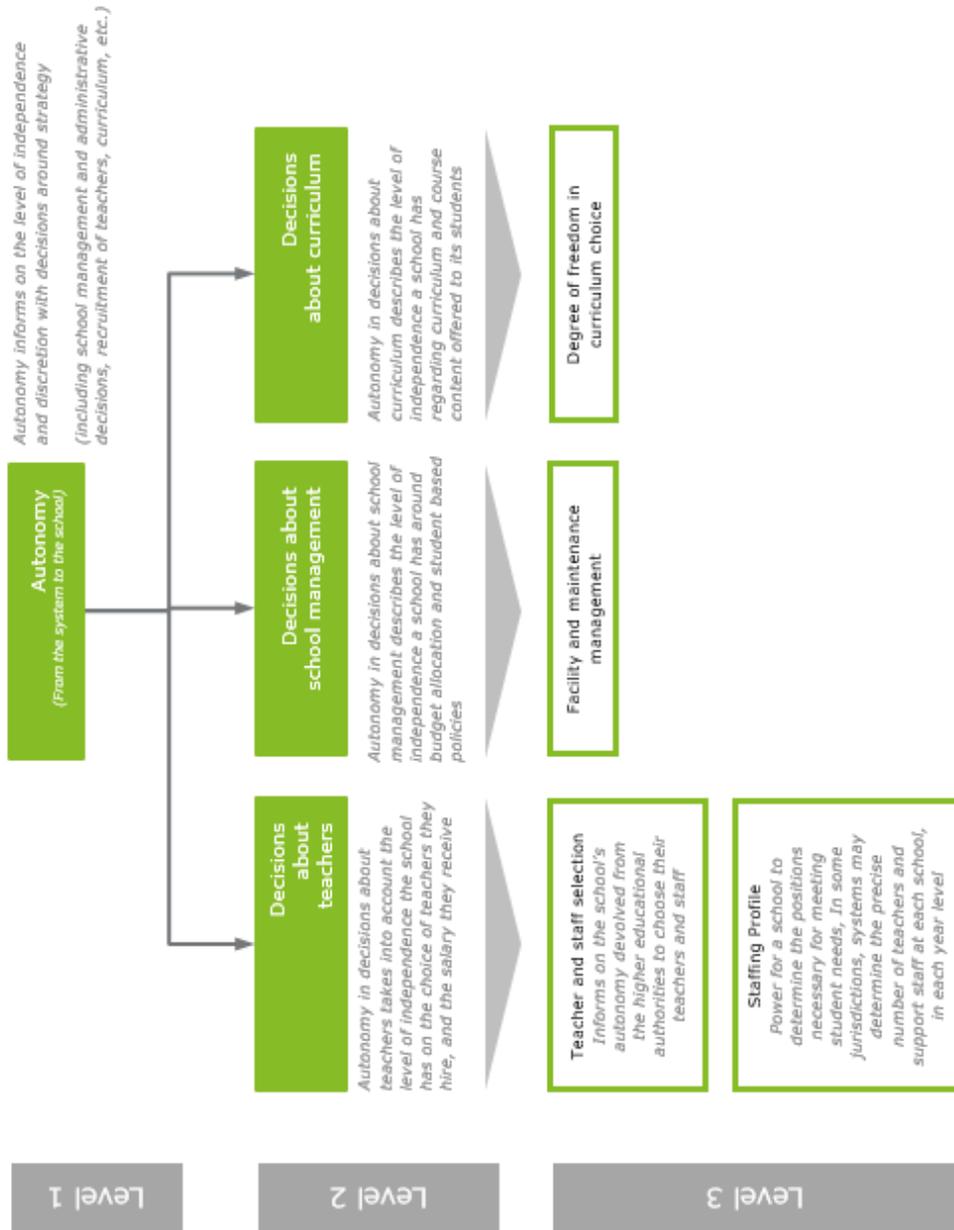
Source: Deloitte Access Economics

Chart A.4: Measurement framework for school leadership, governance and culture



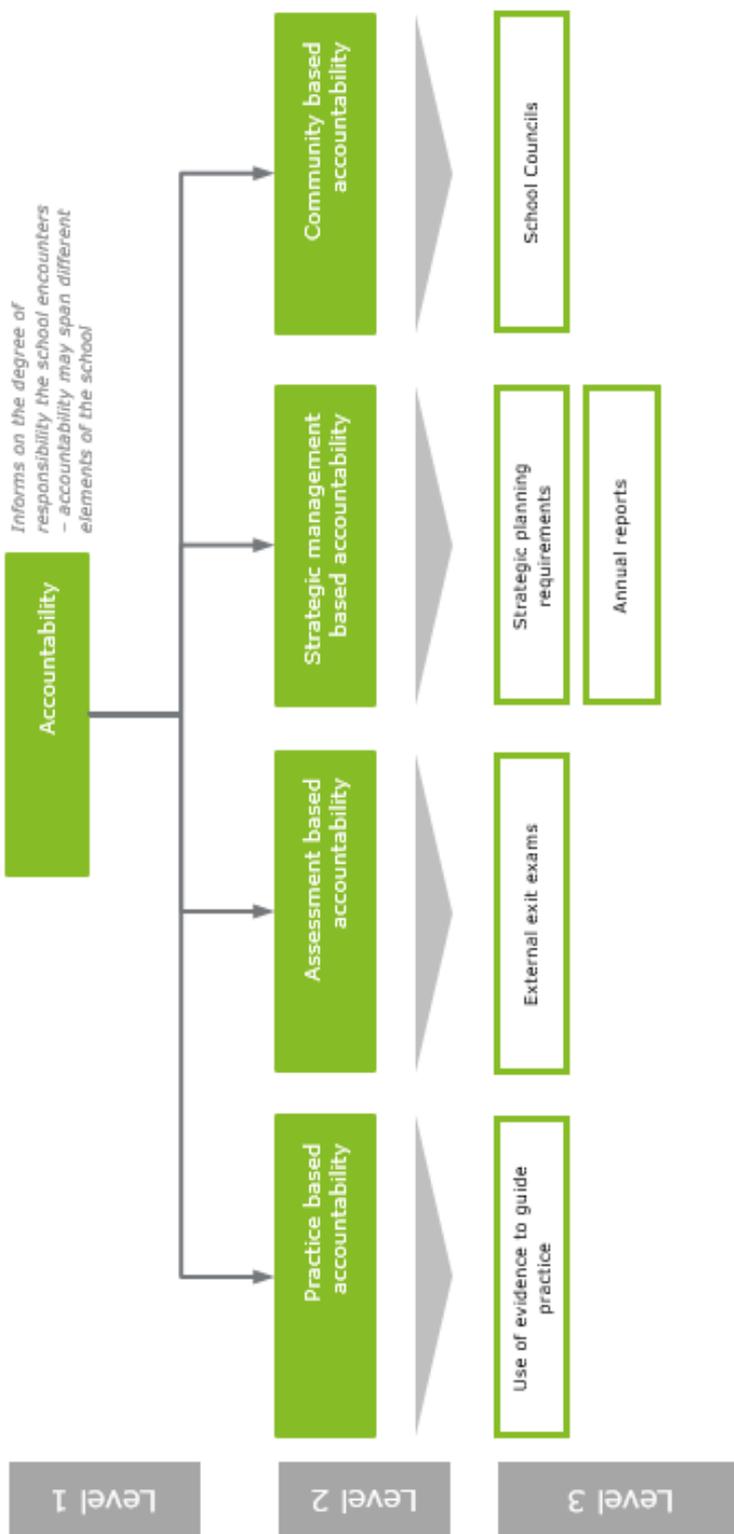
Source: Deloitte Access Economics

Chart A.5: Measurement framework for autonomy



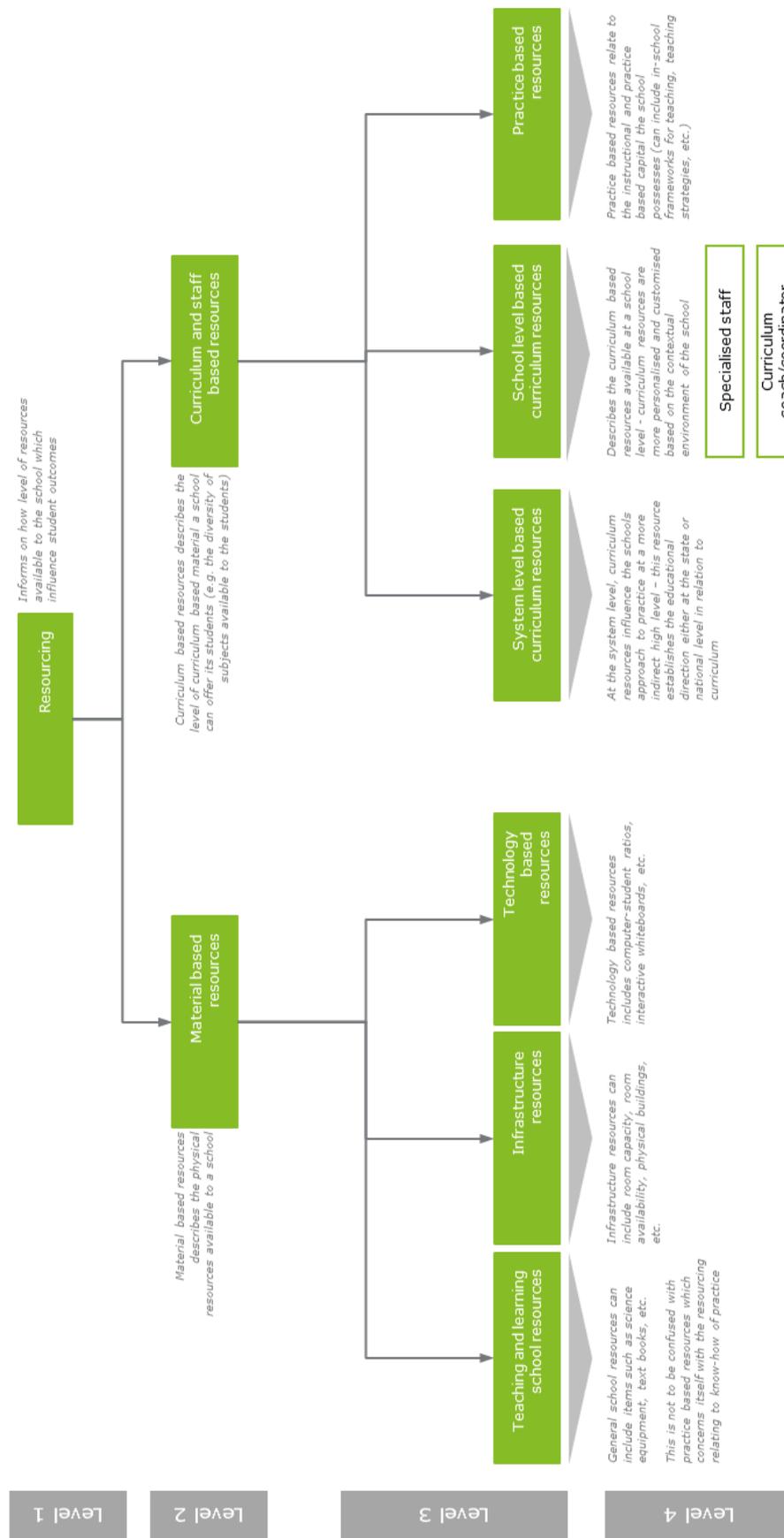
Source: Deloitte Access Economics

Chart A.6: Measurement framework for accountability



Source: Deloitte Access Economics

Chart A.7: Measurement framework for resourcing



Source: Deloitte Access Economics

Appendix B: Evidence for Learning (*Teaching and Learning Toolkit*)

Selected Approaches	Average cost	Evidence security	Months impact	Summary
Arts Participation	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 2 Months	Low impact for low cost, based on moderate evidence.
Aspiration interventions	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 0 Months	Very low impact for moderate cost, based on very limited evidence.
Behavior interventions	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 3 Months	Moderate impact for moderate cost, based on extensive evidence.
Block scheduling	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 0 Months	Very low impact for very low cost, based on limited evidence.
Built environment	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 0 Months	Very low or no impact for low cost, based on very limited evidence.
Collaborative learning	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for very low cost, based on extensive evidence.
Digital Technology	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 4 Months	Moderate impact for moderate cost, based on extensive evidence.
Early years intervention	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for very high cost, based on extensive evidence.
Extending school time	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 2 Months	Low impact for moderate cost, based on moderate evidence.
Feedback	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 8 Months	High impact for moderate cost, based on limited evidence.
Homework (Primary)	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 2 Months	Low impact for very low cost, based on limited evidence.
Homework (Secondary)	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for very low cost, based on moderate evidence.
Individualised instruction	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 2 Months	Low impact for very low cost, based on moderate evidence.
Learning styles	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 2 Months	Low impact for very low cost, based on limited evidence.
Mastery learning	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for very low cost, based on moderate evidence.
Mentoring	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 1 Months	Low impact for moderate cost, based on moderate evidence.
Meta-cognition and self-regulation	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 8 Months	High impact for very low cost, based on extensive evidence.
One to one tuition	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for high cost, based on extensive evidence.
Oral language interventions	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for very low cost, based on extensive evidence.
Outdoor adventure learning	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 4 Months	Moderate impact for moderate cost, based on moderate evidence.
Parental Involvement	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 3 Months	Moderate impact for moderate cost, based on moderate evidence
Peer tutoring	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for very low cost, based on extensive evidence
Performance pay	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 1 Months	Low impact for low cost based on limited evidence.
Phonics	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 4 Months	Moderate impact for very low cost, based on very extensive evidence.
Reading comprehension strategies	\$ - \$ - \$ - \$ - \$ -	★ ★ ★ ★ ★	+ 5 Months	Moderate impact for very low cost, based on extensive evidence.

Selected Approaches	Average cost	Evidence security	Months impact	Summary
Reducing class size	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 3 Months	Moderate impact for high cost, based on moderate evidence.
Repeating a year	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	-4 Months	Negative impact for very high cost, based on moderate evidence.
School uniform	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	0 Months	Very low or no impact for very low cost, based on very limited evidence.
Setting or streaming	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	- 1 Months	Negative impact for very low cost, based on moderate evidence.
Small group tuition	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 4 Months	Moderate impact for moderate cost, based on limited evidence.
Social and emotional learning	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 4 Months	Moderate impact for moderate cost, based on extensive evidence.
Sports participation	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 2 Months	Low impact for moderate cost, based on limited evidence.
Summer schools	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 2 Months	Low impact for moderate cost, based on extensive evidence.
Teaching assistants	\$ \$ \$ \$ \$ \$	★ ★ ★ ★ ★	+ 1 Months	Low impact for high cost, based on limited evidence.

Source: The Education Endowment Foundation (EEF) 2014

Appendix C: Detailed modelling results

C.1. Controls

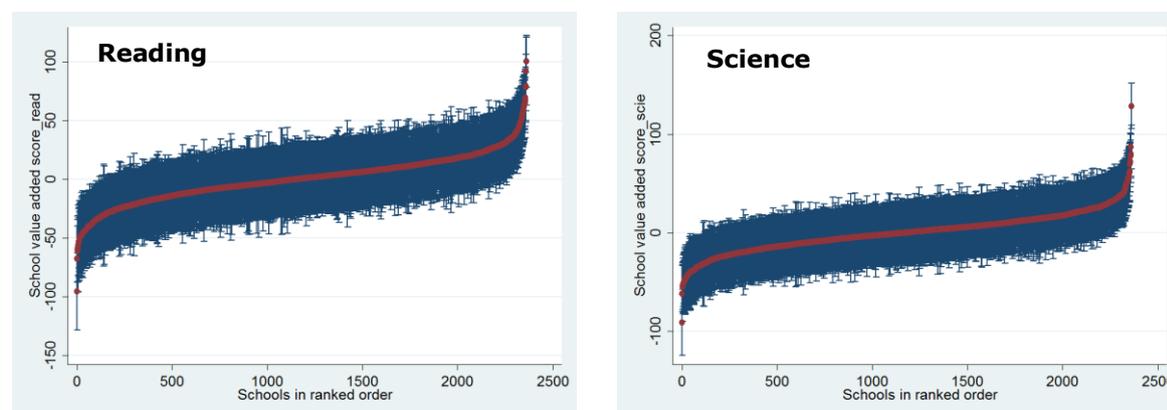
Table C.1: Student and school characteristic controls, TIMSS

Student characteristics	School characteristics	Other characteristics
Home computer	School proportion of economically advantaged students	State ³⁰
Home desk	School proportion of economically disadvantaged students	
Number of books at home	School location	
Age	Total enrolment	
Parental education ³¹	School language	
Gender	School female	
Indigenous status ³²		
Language at home		

Source: Deloitte Access Economics

C.2. Value added results

Chart C.1: Value added by subject (PISA)



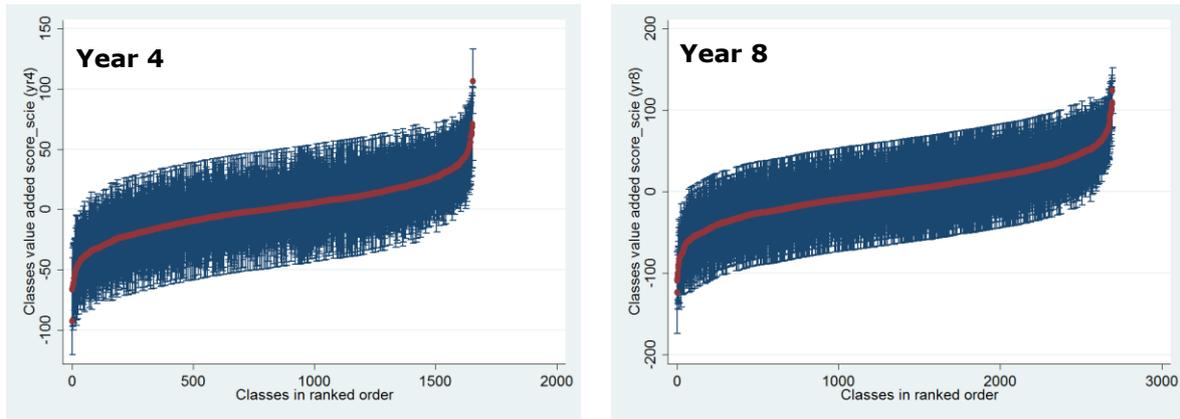
Source: Deloitte Access Economics analysis of PISA data

³⁰ State is available for years 2007, 2011 and 2015 only.

³¹ Question available for year 8 students only.

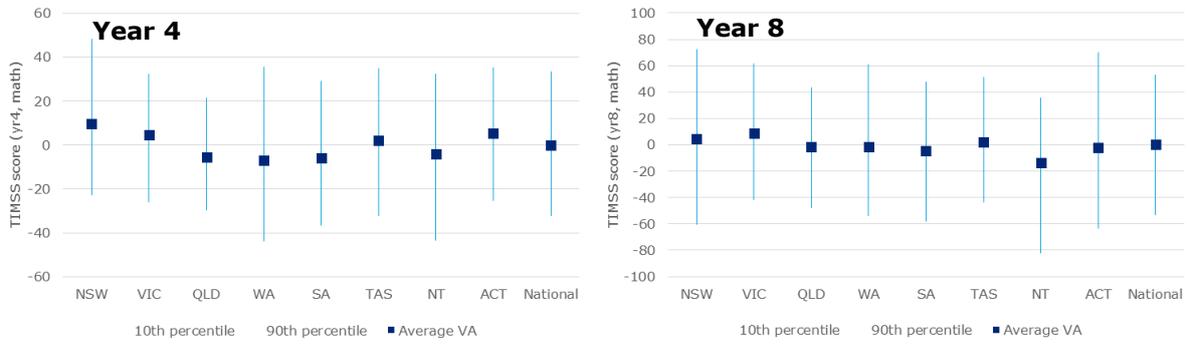
³² Indigenous status available for years 2007, 2011 and 2015 only.

Chart C.2: Value added for science (TIMSS)



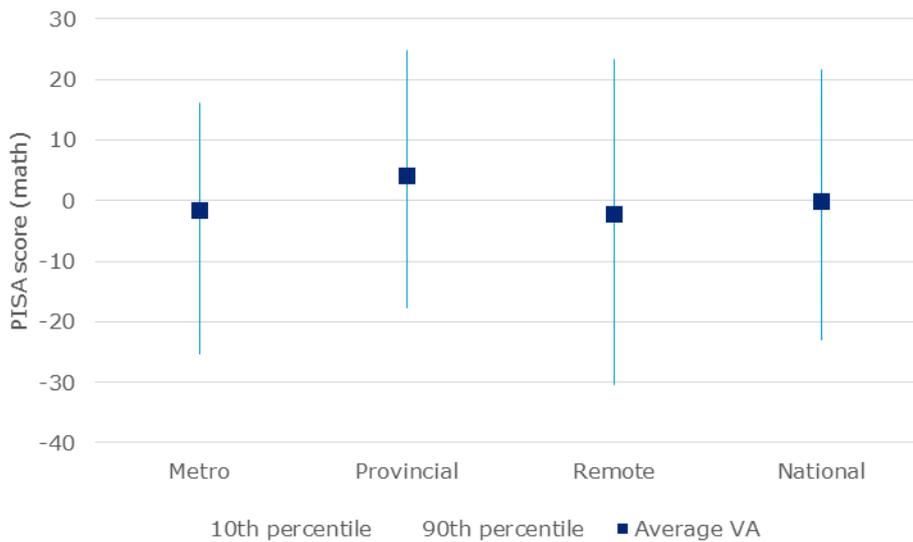
Source: Deloitte Access Economics analysis of TIMSS data

Chart C.3: Distribution in teacher value added to student TIMSS maths scores across states



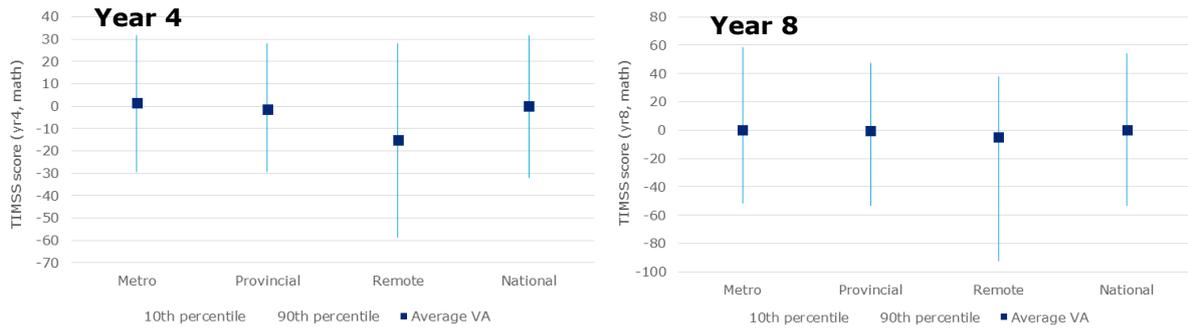
Source: Deloitte Access Economics analysis of TIMSS data

Chart C.4: distribution in school value added to student PISA maths scores across geographies



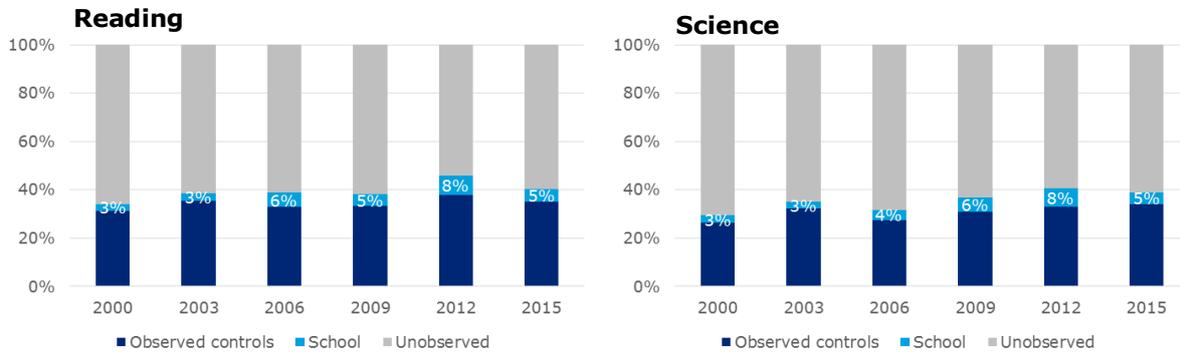
Source: Deloitte Access Economics analysis of PISA data

Chart C.5: distribution of classroom value added to student TIMSS maths scores across geographies



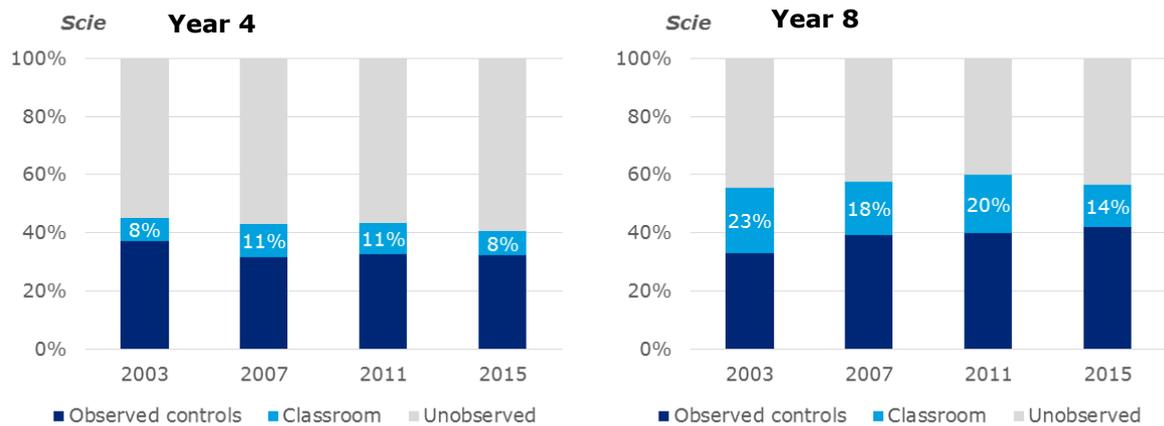
Source: Deloitte Access Economics analysis of TIMSS data

Chart C.6: Accounting for variation in student PISA scores over time, 2000-2015



Source: Deloitte Access Economics analysis of PISA data

Chart C.7: Accounting for variation in student science TIMSS scores over time, 2000-2015



Source: Deloitte Access Economics analysis of TIMSS data

C.3. Drivers of school quality (PISA)

Table C.2: Standardised effect sizes of drivers of school quality on PISA maths scores

Theme	Representative question	Full std effect	Partial std effect
Accountability	Does improvement exist at school? External evaluation (Yes relative to no)	-0.320	0.26
Autonomy - Curriculum - -	Deciding which courses are offered: <Regional or local education authority> ³³ (Yes relative to no)	0.140	1.49
	Establishing student assessment policies: <Regional or local education authority> (Yes relative to no)	-0.260	-1.45
Autonomy - Management - -	Deciding on budget allocations within the school: <Regional or local education authority> (Yes relative to no)	2.53**	2.56
	Deciding on budget allocations within the school: Teachers (Yes relative to no)	-1.660	-1.93
	Formulating the school budget: National education authority (Yes relative to no)	0.640	-1.35
	Establishing student disciplinary policies: <Regional or local education authority> (Yes relative to no)	-4.5***	-1.71
Autonomy - Teacher - -	Establishing teachers' starting salaries: Principal (Yes relative to no)	1.10	3.66
	Selecting teachers for hire: <School governing board> (Yes relative to no)	1.97*	-2.25
	Determining teachers' salary increases: <Regional or local education authority> (Yes relative to no)	-3.2**	-2.48
Classroom organisation and environment - Context - -	Est. percent. <national modal grade for 15-year-olds>. Students with special needs	-1.560	-2.32
Classroom organisation and environment - Environment -	How often does this happen in your <school science> lessons? Students don't listen to what the teacher says. (Less often than average)	4.520	8.34
Class order and cohesion -	How often does this happen in your <school science> lessons? Students cannot work well. (More than average)	-1.530	-8.26

³³ Question: *Regarding your school, who has considerable responsibility for the following tasks?* Responses range from principal and teacher to the governing education agencies

Theme	Representative question	Full std effect	Partial std effect
Classroom organisation and environment - Environment - Peer culture -	To what extent do you disagree or agree about yourself? I am a good listener. (Disagree more than average)	-2.23*	-4.54
	To what extent do you disagree or agree about yourself? I prefer working as part of a team to working alone. (Disagree more than average)	6.53***	8.04
	Extent to which student learning is hindered by: Students intimidating or bullying other students (More than average)	-0.50	-2.08
	To what extent do you disagree or agree about yourself? I enjoy considering different perspectives. (Disagree more than average)	2.97**	3.20
Classroom organisation and environment - Organisation - Practice driven organisation -	School's policy\for students in <national modal grade for 15-year-olds>? Students\group ability within\classes. (More subjects than average)	-1.690	1.21
Classroom organisation and environment - Organisation - Teacher working conditions -	Teachers in TOTAL: Full-time	-11.59**	5.22
	School instruction hindered by: Inadequate or poorly qualified assisting staff. (More than average)	-1.610	-0.54
Resourcing - Material based - Infrastructure -	School instruction hindered by: Inadequate\poor quality phys infrastructure (building, grounds, heating\cooling). (More than average)	3.560	0.78
Resourcing - Material based - Teaching and learning -	If we ever have some extra funding, a big share goes into improvement of our <school science> teaching.(Yes relative to no)	0.560	2.15
	School instruction hindered by: Inadequate or poor quality educational material (e.g. textbooks, IT equipment). (More than average)	-3.37**	-1.00
School leadership, governance, and culture - Parental culture and involvement - -	Did your school implement any measures in: Parental engagement in school (No relative to yes)	-1.410	-0.62
	School provides information\ideas to families about help at home homework and other curriculum-related activities .(Yes relative to no)	0.820	1.60
	School designs effective forms of school-to-home\home-to-school communications\school program\child's progress. .(Yes relative to no)	0.110	1.46

Theme	Representative question	Full std effect	Partial std effect
	<the last academic year>, what proportion. of parents participated in school-related activities? On initiative of child's teachers	1.740	2.46
	There is a <national, state, or district legislation> on including parents in school activities. .(Yes relative to no)	-4.71***	-1.21
	Does improvement exist at school? Consultation aimed at school improvement\experts over a period of six months	0.710	0.59
School leadership, governance, and culture - Policy, mission and goals - Development and collaboration -	Internal evaluation \ Self-evaluation	2.94*	-0.50
	Frequency of <the last academic year>. I use student performance results to develop the school's educational goal	-10	1.47
	We used the data to plan specific action for school development. (Yes, mandatory relative to Yes, school initiative)	-2.29*	-3.45
	Does improvement exist at school? Systematic recording of data such as attendance and professional development (No relative to Yes, mandatory)	0.40	0.27
School leadership, governance, and culture - Policy, mission and goals - Transparency and clarity -	Does improvement exist at school? Written specification of student performance standards.(No relative to Yes, mandatory)	0.460	0.64
	Achievement data used in any of the following <accountability procedures>? Achievement data\provided to parents (Yes relative to no)	-1.30	1.58
	Achievement data used in any of the following <accountability procedures>? Achievement data are posted publicly (Yes, relative to no)	-0.260	-1.40
School leadership, governance, and culture - Principal attributes, culture and integration - Role -	Extent to which student learning is hindered by: Teachers not meeting individual students' needs (More than average)	-2.190	-2.25
	Did your school implement any measures in: Implementation of the curriculum (No relative to Yes)	-1.740	-0.40
	Did your school implement any measures in: Educational staff (e.g. workload, personal	-1.230	-0.76

Theme	Representative question	Full std effect	Partial std effect
	requirement) (No, for other reasons relative to No, because satisfactory)		
	Does improvement exist at school? Teacher mentoring (Yes, mandatory relative to Yes, school initiative)	-1.90	-0.37
	Does improvement exist at school? Seeking written feedback from students (e.g. regarding lessons, teachers, resources (Yes, mandatory relative to Yes, school initiative)	0.270	0.30
	Frequency of <the last academic year>. When a teacher has problems in his\her classroom, I take the initiative to (Less often than average)	1.520	1.97
	Frequency of <the last academic year>. I praise teachers whose students are actively participating in learning. (Less often than average)	-2.480	-1.19
	We put measures derived from the results of external evaluations into practice promptly. (Yes relative to no)	-0.430	-2.54
	Did your school implement any measures in: Student achievement (No, for other reasons relative to Yes)	1.080	-0.16
	Did your school implement any measures in: Teacher professional development (No, because satisfactory relative to Yes)	-0.590	-0.14
School leadership, governance, and culture - Staff culture and involvement - -	Frequency of <the last academic year>. I engage teachers to help build a school culture of continuous improvement. (Less often than average)	-1.80	1.10
Teaching efficacy - Attributes - Qualifications -	Teachers with an <ISCED Level 5A Bachelor degree> qualification: Full-time	6.980	5.54
	Teachers with an <ISCED Level 6> qualification: Part-time	-0.350	1.49
Teaching efficacy - Practice - Approach to teaching and learning - Assessment	Are teacher-developed tests used in school? To award certification to students (Yes relative to no)	2.33*	1.59
	Are <standardized tests> used in school? To award certification to students (Yes relative to no)	1.80	2.91

Theme	Representative question	Full std effect	Partial std effect
	Are teacher-developed tests used in school? To identify aspects of instruction or curriculum that should be improved (Yes relative to no)	1.470	1.68
	Are <standardized tests> used in school? To group students for instructional purposes (Yes relative to no)	0.630	1.48
	Are teacher-developed tests used in school? To adapt teaching to the students' needs (Yes relative to no)	-2.020	1.15
	Are teacher-developed tests used in school? To inform parents about child's progress (Yes relative to no)	-0.650	1.38
Teaching efficacy - Practice - Approach to teaching and learning - Instructional approaches	When learning <school science>? Students are allowed to design their own experiments. (More often than average)	-2.45**	-8.47
Teaching efficacy - Practice - Approach to teaching and learning - Targeted teaching strategies	How often does this happen in <school science>? The teacher provides individual help when a student has difficulties (More often than average)	1.93**	7.94
	Teachers disciplined me more harshly than other students. (More often than average)	-3.8*	-9.53
	Teachers graded me harder than they graded other students. (More often than average)	-0.940	-5.50
Teaching efficacy - Practice - Wellbeing and development - Professionalism	Teachers called on me less often than they called on other students. (More often than average)	-2.75*	-6.79
	Teachers said something insulting to me in front of others. (More often than average)	-2.85*	-9.17
	Teachers ridiculed me in front of others. (More often than average)	0.890	-8.30
Teaching efficacy - Practice - Wellbeing and development - Support and development	Teachers in our school cooperate by exchanging ideas or material when teaching specific units or series of lessons. (Yes relative to no)	5.56**	1.52

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of PISA data

Table C.3: Standardised effect sizes of drivers of school quality on PISA science scores

Theme	Representative question	Full std effect	Partial std effect
Accountability	Does improvement exist at school? External evaluation (Yes, school initiative relative to Yes, mandatory)	-0.67**	0.26
Autonomy - Curriculum - -	Deciding which courses are offered: <Regional or local education authority> ³⁴ (Yes relative to no)	0.30	1.49
	Establishing student assessment policies: <Regional or local education authority> (Yes relative to no)	0.220	-1.45
Autonomy - Management - -	Deciding on budget allocations within the school: <Regional or local education authority> (Yes relative to no)	1.480	2.56
	Deciding on budget allocations within the school: Teachers (Yes relative to no)	-3.3***	-1.93
	Formulating the school budget: National education authority (Yes relative to no)	0.640	-1.35
	Establishing student disciplinary policies: <Regional or local education authority> (Yes relative to no)	-1.540	-1.71
Autonomy - Teacher - -	Establishing teachers' starting salaries: Principal (Yes relative to no)	-0.80	3.66
	Selecting teachers for hire: <School governing board>(Yes relative to no)	3.33***	-2.25
	Determining teachers' salary increases: <Regional or local education authority> (Yes relative to no)	-3.17*	-2.48
Classroom organisation and environment - Context - -	Est. percent. <national modal grade for 15-year-olds>. Students with special needs	-2.010	-2.32
Classroom organisation and environment - Environment -	How often does this happen in your <school science> lessons? Students don't listen to what the teacher says. (Less often than average)	5.880	8.34
Class order and cohesion -	How often does this happen in your <school science> lessons? Students cannot work well. (More often than average)	-2.07*	-8.26

³⁴ Question: *Regarding your school, who has considerable responsibility for the following tasks?* Responses range from principal and teacher to the governing education agencies

Theme	Representative question	Full std effect	Partial std effect
Classroom organisation and environment - Environment - Peer culture -	To what extent do you disagree or agree about yourself? I am a good listener. (Disagree more than average)	-3.65*	-4.54
	To what extent do you disagree or agree about yourself? I prefer working as part of a team to working alone. (Disagree more than average)	8.97***	8.04
	Extent to which student learning is hindered by: Students intimidating or bullying other students (More than average)	-1.940	-2.08
	To what extent do you disagree or agree about yourself? I enjoy considering different perspectives. (Agree more than average)	4.81*	3.20
Classroom organisation and environment - Organisation - Practice driven organisation -	School's policy\for students in <national modal grade for 15-year-olds>? Students\group ability within\classes. (More often than average)	-1.050	-1.21
Classroom organisation and environment - Organisation - Teacher working conditions -	Teachers in TOTAL: Full-time	-5.10	5.22
	School instruction hindered by: Inadequate or poorly qualified assisting staff.(More than average)	2.840	-0.54
Resourcing - Material based - Infrastructure -	School instruction hindered by: Inadequate\poor quality phys infrastructure (building, grounds, heating\cooling). (More than average)	-2.250	-0.78
Resourcing - Material based - Teaching and learning -	If we ever have some extra funding, a big share goes into improvement of our <school science> teaching. (Yes relative to no)	0.370	2.15
	School instruction hindered by: Inadequate or poor quality educational material (e.g. textbooks, IT equipment). (More than average)	-1.280	1.00
School leadership, governance, and culture - Parental culture and involvement - -	Did your school implement any measures in: Parental engagement in school (No relative to yes)	-0.970	-0.62
	School provides information\ideas to families about help at home homework and other curriculum-related activities (Yes relative to no)	-0.460	1.60
	School designs effective forms of school-to-home\home-to-school communications\school program\child's progress. (Yes relative to no)	-0.310	1.46

Theme	Representative question	Full std effect	Partial std effect
	<the last academic year>, what proportion of parents participated in school-related activities? On initiative of child's teachers	1.730	2.46
	There is a <national, state, or district legislation> on including parents in school activities. (Yes relative to no)	-4.49***	-1.21
	Does improvement exist at school? Consultation aimed at school improvement\experts over a period of six months (Yes, mandatory relative to Yes, school initiative)	-2.060	-0.59
School leadership, governance, and culture - Policy, mission and goals - Development and collaboration -	Internal evaluation \ Self-evaluation (No relative to Yes, school initiative)	0.840	0.50
	Frequency of <the last academic year>. I use student performance results to develop the school's educational goal (Less often than average)	-2.340	-1.47
	We used the data to plan specific action for school development. (Yes relative to no)	-4.55***	-3.45
	Does improvement exist at school? Systematic recording of data such as attendance and professional development (Yes, school initiative relative to Yes, mandatory)	-1.640	0.27
School leadership, governance, and culture - Policy, mission and goals - Transparency and clarity -	Does improvement exist at school? Written specification of student performance standards (Yes, school initiative relative to Yes, mandatory)	2.350	0.64
	Achievement data used in any of the following <accountability procedures>? Achievement data\provided to parents (Yes relative to no)	-1.650	1.58
	Achievement data used in any of the following <accountability procedures>? Achievement data are posted publicly (Yes relative to no)	0.760	-1.40
School leadership, governance, and culture - Principal attributes, culture and integration - Role -	Extent to which student learning is hindered by: Teachers not meeting individual students' needs (More often than average)	1.150	-2.25
	Did your school implement any measures in: Implementation of the curriculum (No relative to Yes)	0.640	-0.40
	Did your school implement any measures in: Educational staff (e.g. workload, personal	0.870	-0.76

Theme	Representative question	Full std effect	Partial std effect
	requirement) (Yes relative to No because satisfactory)		
	Does improvement exist at school? Teacher mentoring (Yes, mandatory relative to Yes, school initiative)	-2.620	-0.37
	Does improvement exist at school? Seeking written feedback from students (e.g. regarding lessons, teachers, resources (No relative to Yes, school initiative)	-0.970	-0.30
	Frequency of <the last academic year>. When a teacher has problems in his\her classroom, I take the initiative to (More often than average)	-1.940	1.97
	Frequency of <the last academic year>. I praise teachers whose students are actively participating in learning. (Less often than average)	2.170	1.19
	We put measures derived from the results of external evaluations into practice promptly. (Yes relative to no)	3.63***	-2.54
	Did your school implement any measures in: Student achievement (No, relative to yes)	-0.50	-0.16
	Did your school implement any measures in: Teacher professional development (No, relative to yes)	0.390	-0.14
School leadership, governance, and culture - Staff culture and involvement - -	Frequency of <the last academic year>. I engage teachers to help build a school culture of continuous improvement. (More often than average)	2.530	1.10
Teaching efficacy - Attributes - Qualifications -	Teachers with an <ISCED Level 5A Bachelor degree> qualification: Full-time	-4.490	5.54
	Teachers with an <ISCED Level 6> qualification: Part-time	-0.850	1.49
Teaching efficacy - Practice - Approach to teaching and learning - Assessment	Are teacher-developed tests used in school? To award certification to students (Yes relative to no)	1.250	1.59
	Are <standardized tests> used in school? To award certification to students (Yes relative to no)	1.250	1.59

Theme	Representative question	Full std effect	Partial std effect
	Are teacher-developed tests used in school? To identify aspects of instruction or curriculum that should be improved (Yes relative to no)	3.95**	1.68
	Are <standardized tests> used in school? To group students for instructional purposes (Yes relative to no)	-0.870	1.48
	Are teacher-developed tests used in school? To adapt teaching to the students' needs (Yes relative to no)	-3.72**	1.15
	Are teacher-developed tests used in school? To inform parents about child's progress (Yes relative to no)	2.130	1.38
Teaching efficacy - Practice - Approach to teaching and learning - Instructional approaches	When learning <school science>? Students are allowed to design their own experiments. (More often than average)	-2.94**	-8.47
Teaching efficacy - Practice - Approach to teaching and learning - Targeted teaching strategies	How often does this happen in <school science>? The teacher provides individual help when a student has difficulties (Less often than average)	-3.19**	-7.94
	Teachers disciplined me more harshly than other students. (More often than average)	-5.17*	-9.53
	Teachers graded me harder than they graded other students. (More often than average)	-1.560	-5.50
Teaching efficacy - Practice - Wellbeing and development - Professionalism	Teachers called on me less often than they called on other students. (More often than average)	-1.910	-6.79
	Teachers said something insulting to me in front of others. (More often than average)	-2.530	-9.17
	Teachers ridiculed me in front of others. (More often than average)	-0.590	-8.30
Teaching efficacy - Practice - Wellbeing and development - Support and development	Teachers in our school cooperate by exchanging ideas or material when teaching specific units or series of lessons. (Yes relative to no)	-0.010	1.52

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of PISA data

Table C.4: Standardised effect sizes of drivers of school quality on TIMSS year 4 maths scores

Theme	Representative question	Full std effect	Partial std effect
Classroom organisation and environment - Environment - Order and cohesion -	GEN\THINKING ABT CURR SCH\STUD BEHAVE (Agree less than average)	-1.280	-7.53
	GEN\LIMIT TEACHING\MENTAL DISABILITIES ³⁵ (More than average)	-0.680	-4.16
	GEN\LIMIT TEACHING\PHYSICAL DISABILITIES ³⁶ (More than average)	-2.38**	-3.08
Classroom organisation and environment - Organisation - Compositional strategies of classroom organisation -	GEN\NUMBER OF STUDENTS IN THE CLASS	7.24***	4.18
Classroom organisation and environment - Organisation - Practice driven organisation -	GEN\STUDENT ACHIEVEMENT USED\SCIENCE (Yes relative to no)	7.12***	2.66
Classroom organisation and environment - Organisation - Teacher working conditions -	GEN\AGREEMENT\NEED MORE TIME TO ASSIST (Yes relative to no)	-1.470	-2.28
	GEN\AGREEMENT\TOO MANY HOURS (Agree more than average)	2.690	3.24
	GEN\LANGUAGE DIFFICULTIES (Yes relative to no)	-4.410	-4.22
Resourcing - Curriculum and staff based - School based -	GEN\SHORTAGE\MAT\TEACH SPEC MATH (Agree more than average)	-4.32**	-1.59
Resourcing - Material based - Infrastructure -	GEN\SEVERITY PROBLEM\MAINTENANCE WORK (Yes relative to no)	-2.070	2.15
	GEN\SHORTAGE\GEN\INSTRUCTIONAL SPACE (Agree more than average)	-1.630	-2.18
Resourcing - Material based - Teaching and learning -	GEN\SHORTAGE\SCI\LIBRARY RESOURCES (Agree more than average)	-6.53*	-3.06

³⁵ Question: *In your view, to what extent do the following (students with mental disabilities) limit how you teach this class?*

This question informs on the extent to which students with mental disabilities influence class order and cohesion and the resultant influence on student outcomes

³⁶ Question: *In your view, to what extent do the following (students with physical disabilities) limit how you teach this class?*

This question informs on the extent to which students with physical disabilities influence class order and cohesion and the resultant influence on student outcomes

Theme	Representative question	Full std effect	Partial std effect
	GEN\SHORTAGE\MAT\LIBRARY RESOURCES (Yes relative to no)	-5.990	2.62
School leadership, governance and culture - Parental culture and involvement - -	GEN\CHARACTERIZE\PARENTAL EXPECTATIONS (Yes relative to no)	-3.940	-5.41
School leadership, governance and culture - Policy, mission and goals - Development and collaboration -	GEN\CHARACTERIZE\COLLABORATION TO PLAN (Higher than average)	-1.930	2.39
	SCI\SCIENCE TAUGHT AS SEPARATE SUBJECT (No relative to yes)	-7.16*	-2.49
	GEN\FREE MEALS\BREAKFAST (Yes relative to no)	-3.40	-6.89
School leadership, governance and culture - Policy, mission and goals - School level approach to practice -	GEN\EMPHASIS\WASHING HANDS (Yes relative to no)	-3.45*	-3.19
	GEN\INSTRUCTIONAL DAYS PER YEAR (Yes relative to no)	4.59**	2.22
	GEN\ASSIST WITH SCHOOLWORK (Lower than average)	-0.940	-0.18
School leadership, governance and culture - Policy, mission and goals - Transparency and clarity -	GEN\CHARACTERIZE\CLARITY OF OBJECTIVES	2.560	-2.45
School leadership, governance and culture - Principal attributes, culture and integration - Attributes -	GEN\HIGHEST LEVEL OF FORMAL EDUCATION (Lower than average)	-0.30	-3.88
School leadership, governance and culture - Staff culture and involvement - -	GEN\SCH CHARACTER\TCH UNDERSTANDING (Lower than average)	-0.660	-3.77
	GEN\CHARACTERIZE\LEADERSHIP SUPPORT (Higher than average)	4.03*	2.91
	GEN\<SPECIALIZATION>\OTHER SUBJECT (Yes relative to no)	-1.710	-2.11
Teaching efficacy - Attributes - Qualifications - Type of teacher education	GEN\<SPECIALIZATION>\MATH (Yes relative to no)	1.66*	4.13
	GEN\MAJOR AREA OF STUDY\SCI (Yes relative to no)	0.730	2.71

Theme	Representative question	Full std effect	Partial std effect
	GEN\ <specialization>\language-reading (yes="" no)<="" relative="" td="" to=""> <td>1.270</td> <td>1.94</td> </specialization>\language-reading>	1.270	1.94
Teaching efficacy - Practice - Approach to teaching and learning - Assessment	SCI\EMPHASIS\EVALUATION OF WORK (Yes relative to no)	-0.320	1.56
Teaching efficacy - Practice - Approach to teaching and learning - Curriculum	MAT\TOPICS\GEOMETRY\ANGLES (Mostly taught the year before, relative to Mostly taught this year)	0.620	-2.93
	MAT\TOPICS\NUMBER\CONCEPTS OF DECIMALS (Not yet taught, relative to Mostly taught this year)	-4.97**	-3.87
Teaching efficacy - Practice - Approach to teaching and learning - Instructional approaches	MAT\PERMITTED TO USE CALCULATORS (Yes, unrestricted relative to Yes, restricted)	-0.250	-2.60
	MAT\ASK STUDENTS\WORK WHILE OCCUPIED (More often than average)	-3.210	-3.54
	GEN\HOW OFTEN\EXPRESS IDEAS (Yes relative to no)	-0.80	-2.35
Teaching efficacy - Practice - Approach to teaching and learning - Lesson planning and collaboration	GEN\INTERACTIONS TEACHERS\SHARE LEARNING (Yes relative to no)	-3.660	-3.12
Teaching efficacy - Practice - Wellbeing and development - Engagement and wellbeing	SCI\CONFIDENT\MAKE SCIENCE RELEVANT (Lower than average)	-1.430	-3.69
	SCI\PREPARED\LIFE\RELATIONSHIPS IN COMM (Lower than average)	0.870	-2.25
	MAT\PREPARED\NUMBER\NUMBER PATTERNS (Lower than average)	1.76***	-1.70
	MAT\PREPARED\NUMBER\CONCEPTS OF DECIMALS (Lower than average)	-3.5***	-3.61
	MAT\PREPARED\NUMBER\USING FRACTIONS (Lower than average)	-0.840	-3.71
Teaching efficacy - Practice - Wellbeing and development - Professionalism	GEN\DEGREE PROBS TEACH\ABSENTEEISM (More than average)	-2.390	-2.82
	GEN\DEGREE PROBS TEACH\ARRIVING LATE AT SCHOOL (More than average)	-0.380	-2.95

Theme	Representative question	Full std effect	Partial std effect
	MAT\PROF DEVELOPMENT\MATH PEDAGOGY (Less than average)	-0.320	-1.90

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of TIMSS data

Table C.5: Standardised effect sizes of drivers of school quality on TIMSS year 4 science scores

Theme	Representative question	Full std effect	Partial std effect
	GEN\THINKING ABT CURR SCH\STUD BEHAVE (Agree less than average)	-2.390	-7.53
Classroom organisation and environment - Environment - Order and cohesion -	GEN\LIMIT TEACHING\MENTAL DISABILITIES ³⁷ (Agree more than average)	-0.750	-4.16
	GEN\LIMIT TEACHING\PHYSICAL DISABILITIES ³⁸ (Agree more than average)	-1.520	3.08
Classroom organisation and environment - Organisation - Compositional strategies of classroom organisation -	GEN\NUMBER OF STUDENTS IN THE CLASS	7.47***	4.18
Classroom organisation and environment - Organisation - Practice driven organisation -	GEN\STUDENT ACHIEVEMENT USED\SCIENCE (Yes relative to no)	6.06***	2.66
Classroom organisation and environment - Organisation - Teacher working conditions -	GEN\AGREEMENT\NEED MORE TIME TO ASSIST (Yes relative to no)	-0.770	-2.28
	GEN\AGREEMENT\TOO MANY HOURS (Agree more than average)	2.280	3.24
	GEN\LANGUAGE DIFFICULTIES (Yes relative to no)	-1.460	-4.22
Resourcing - Curriculum and staff based - School based -	GEN\SHORTAGE\MAT\TEACH SPEC MATH (Agree more than average)	-3.52*	-1.59

³⁷ Question: *In your view, to what extent do the following (students with mental disabilities) limit how you teach this class?*

This question informs on the extent to which students with mental disabilities influence class order and cohesion and the resultant influence on student outcomes

³⁸ Question: *In your view, to what extent do the following (students with physical disabilities) limit how you teach this class?*

This question informs on the extent to which students with physical disabilities influence class order and cohesion and the resultant influence on student outcomes

Theme	Representative question	Full std effect	Partial std effect
Resourcing - Material based - Infrastructure -	GEN\SEVERITY PROBLEM\MAITENANCE WORK (Yes relative to no)	-2.10	-2.15
	GEN\SHORTAGE\GEN\INSTRUCTIONAL SPACE (More than average)	-0.760	-2.18
Resourcing - Material based - Teaching and learning -	GEN\SHORTAGE\SCI\LIBRARY RESOURCES (Less than average)	-7.43***	-3.06
	GEN\SHORTAGE\MAT\LIBRARY RESOURCES (Less than average)	-5.12*	2.62
School leadership, governance and culture - Parental culture and involvement - -	GEN\CHARACTERIZE\PARENTAL EXPECTATIONS (Yes relative to no)	1.920	-5.41
School leadership, governance and culture - Policy, mission and goals - Development and collaboration -	GEN\CHARACTERIZE\COLLABORATION TO PLAN (Lower than average)	-1.220	-2.39
School leadership, governance and culture - Policy, mission and goals - School level approach to practice -	SCI\SCIENCE TAUGHT AS SEPARATE SUBJECT (No relative to yes)	-2.470	-2.49
	GEN\FREE MEALS\BREAKFAST (Yes relative to no)	-3.370	-6.89
	GEN\EMPHASIS\WASHING HANDS (Yes relative to no)	-2.080	-3.19
	GEN\INSTRUCTIONAL DAYS PER YEAR	0.890	2.22
	GEN\ASSIST WITH SCHOOLWORK (Yes relative to no)	1.030	0.18
School leadership, governance and culture - Policy, mission and goals - Transparency and clarity -	GEN\CHARACTERIZE\CLARITY OF OBJECTIVES (Lower than average)	2.98*	-2.45
School leadership, governance and culture - Principal attributes, culture and integration - Attributes -	GEN\HIGHEST LEVEL OF FORMAL EDUCATION (Higher than average)	1.210	3.88
School leadership, governance and culture - Staff culture and involvement - -	GEN\SCH CHARACTER\TCH UNDERSTANDING (Lower than average)	-1.440	-3.77
	GEN\CHARACTERIZE\LEADERSHIP SUPPORT (Yes relative to no)	2.630	2.91
	GEN\ <specialization>>\OTHER SUBJECT (Yes relative to no)</specialization>	-0.540	-2.11

Theme	Representative question	Full std effect	Partial std effect
Teaching efficacy - Attributes - Qualifications - Type of teacher education	GEN\ <specialization>\MATH (Yes relative to no)</specialization>	0.890	4.13
	GEN\MAJOR AREA OF STUDY\SCI (Yes relative to no)	-0.160	2.71
	GEN\ <specialization>\LANGUAGE-READING (Yes relative to no)</specialization>	-0.070	1.94
Teaching efficacy - Practice - Approach to teaching and learning - Assessment	SCI\EMPHASIS\EVALUATION OF WORK	-0.420	1.56
	SCI\TOPICS\EARTH\WEATHER CONDITIONS (Not yet taught, relative to Mostly taught this year)	-2.44*	-2.23
	SCI\TOPICS\PHYSICAL\CHANGES IN MATERIALS (Mostly taught this year relative to not yet taught)	1.730	1.70
	SCI\TOPICS\PHYSICAL\FORMING MIXTURES (Not yet taught, relative to Mostly taught this year)	-1.90	-2.29
	GEN\HOW OFTEN\EXPRESS IDEAS (Yes relative to no)	0.750	-2.35
	SCI\ASK STUDENTS\PLAN EXPERIMENTS (Less than average)	-1.620	-3.28
	SCI\ASK STUDENTS\DEMONSTRATE EXPERIMENT (Less than average)	-2.7**	1.72
	SCI\ASK STUDENTS\EXPLAIN NEW CONTENT(Less than average)	-1.340	-2.46
Teaching efficacy - Practice - Approach to teaching and learning - Lesson planning and collaboration	GEN\INTERACTIONS TEACHERS\SHARE LEARNING (Less than average)	-1.70	-3.12
Teaching efficacy - Practice - Wellbeing and development - Engagement and wellbeing	SCI\CONFIDENT\MAKE SCIENCE RELEVANT (Less than average)	0.850	-3.69
	SCI\PREPARED\LIFE\RELATIONSHIPS IN COMM (Less than average)	0.830	-2.25
	GEN\DEGREE PROBS TEACH\ABSENTEEISM (More than average)	-1.940	-2.82

Theme	Representative question	Full std effect	Partial std effect
Teaching efficacy - Practice - Wellbeing and development - Professionalism	GEN\DEGREE PROBS TEACH\ARRIVING LATE AT SCHOOL (More than average)	10	-2.95
Teaching efficacy - Practice - Wellbeing and development - Support and development	SCI\PROF DEVELOPMENT\STUDENT NEEDS (Yes relative to no)	0.30	-1.73
	SCI\PROF DEVELOPMENT\SCIENCE PEDAGOGY (More than average)	0.390	1.45

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of TIMSS data

Table C.6: Standardised effect sizes of drivers of school quality on TIMSS year 8 maths scores

Theme	Representative question	Full std effect (absolute value)	Partial std effect (absolute value)
Classroom organisation and environment - Environment - Order and cohesion -	GEN\THINKING ABT CURR SCH\STUD BEHAVE (Disagree more than average)	-9.25***	-20.40
	GEN\LIMIT TEACHING\PHYSICAL DISABILITIES (More than average)	-1.98**	-14.02
	GEN\LIMIT TEACHING\UNINTERESTED STUDENTS (More than average)	6.28***	18.06
Classroom organisation and environment - Organisation - Compositional strategies of classroom organisation -	GEN\NUMBER OF STUDENTS IN THE CLASS	12.09***	17.79
Classroom organisation and environment - Organisation - Practice driven organisation -	GEN\STUDENT ACHIEVEMENT USED TO ASSIGN\MATH (Yes relative to no)	5.06*	3.87
Classroom organisation and environment - Organisation - Teacher working conditions	GEN\LANGUAGE DIFFICULTIES (Yes relative to no)	-6.35***	-5.36
	GEN\AGREEMENT\CHANGES IN CURRICULUM (Agree more than average)	-4.890	3.31
	GEN\AGREEMENT\NEED MORE TIME TO ASSIST (Disagree more than average)	-1.440	4.75
Resourcing - Curriculum and staff based - School based -	GEN\SHORTAGE\MATH\TEACH SPEC MATH (More than average)	-4.62*	-12.42

Theme	Representative question	Full std effect (absolute value)	Partial std effect (absolute value)
	GEN\SEVERITY PROBLEM\BUILDING REPAIR (More than average)	3.450	-8.67
	GEN\EXISTING SCIENCE LABORATORY (Yes relative to no)	15.02***	4.05
	GEN\SEVERITY PROBLEM\MAINTENANCE WORK (More than average)	2.670	-5.98
Resourcing - Material based - Infrastructure	GEN\SEVERITY PROBLEM\CLASSRMS NOT CLEANED (More than average)	2.830	-4.27
	GEN\SHORTAGE\GEN\INSTRUCTIONAL MATERIAL (More than average)	-6.060	-8.79
	GEN\SHORTAGE\SCI\LIBRARY RESOURCES (More than average)	8.350	-8.57
Resourcing - Material based - Teaching and learning	GEN\SHORTAGE\MATH\LIBRARY RESOURCES (More than average)	1.580	-8.18
	GEN\SHORTAGE\GEN\COMP TECHNOLOGY (More than average)	-2.820	-10.40
Resourcing - Material based - Technology -	GEN\SEVERITY PROBLEM\INADEQUATE SUPPORT FOR TECH (More than average)	-2.230	-8.70
School leadership, governance and culture - Parental culture and involvement - -	GEN\SCH CHARACTER\PARENTAL COMMITMENT (Lower than average)	6.520	-10.58
School leadership, governance and culture - Policy, mission and goals - Development and collaboration -	GEN\CHARACTERIZE\COLLABORATION TO PLAN (Higher than average)	-2.910	-6.74
	GEN\USE INCENTIVES\OTHER (Yes relative to no)	-8.66*	-3.76
School leadership, governance and culture - Policy, mission and goals - Hiring and recruitment -	GEN\FILL TEACHING VACANCIES\MATH (More difficult than average)	1.50	-4.81
	GEN\THINKING ABT CURR SCH\CLEAR RULES (Disagree more than average)	-2.390	-11.28
School leadership, governance and culture - Policy, mission and goals - School level approach to practice - -	GEN\FREE MEALS\BREAKFAST (Yes relative to no)	3.630	-15.40

Theme	Representative question	Full std effect (absolute value)	Partial std effect (absolute value)
	GEN\ASSIST WITH SCHOOLWORK (Yes relative to no)	2.730	-7.15
School leadership, governance and culture - Policy, mission and goals - Transparency and clarity -	GEN\CHARACTERIZE\CLARITY OF OBJECTIVES (Lower than average)	4.20	-8.93
	GEN\DEGREES IN EDUCATION LEADERSHIP\ISCED 7	-3.60	3.82
	GEN\YEARS PRINCIPAL ALTOGETHER	-3.810	3.85
School leadership, governance and culture - Principal attributes, culture and integration - Attributes -	GEN\DEGREES IN EDUCATION LEADERSHIP\ISCED 8	11.34***	-2.89
School leadership, governance and culture - Principal attributes, culture and integration - Role -	GEN\CHARACTERIZE\AMOUNT OF INSTR SUPPORT (Lower than average)	0.810	-3.98
School leadership, governance and culture - Staff culture and involvement - -	GEN\SCH CHARACTER\TCH UNDERSTANDING (Higher than average)	6.860	9.19
School leadership, governance and culture - Staff culture and involvement - -	GEN\CHARACTERIZE\TCHS EXPECTATIONS (Higher than average)	5.850	7.07
Teaching efficacy - Attributes - Characteristics - Personal characteristics	GEN\SEX OF TEACHER (Male relative to female)	14.53***	-4.80
	GEN\YEARS BEEN TEACHING	-3.880	4.04
Teaching efficacy - Attributes - Qualifications - Level of teacher education	GEN\LEVEL OF FORMAL EDUCATION COMPLETED (Higher than average)	-4.24**	-7.99
	GEN\MAJOR AREA OF STUDY\OTHER (Yes relative to no)	9.19***	-2.84
	GEN\MAJOR AREA OF STUDY\PHYSICS (Yes relative to no)	-0.30	4.26
	GEN\MAJOR AREA OF STUDY\BIOLOGY (Yes relative to no)	0.250	-3.93
Teaching efficacy - Attributes - Qualifications - Type of teacher education	GEN\MAJOR AREA OF STUDY\MATHEMATICS (Yes relative to no)	18.12***	5.19

Theme	Representative question	Full std effect (absolute value)	Partial std effect (absolute value)
	GEN\MAJOR AREA OF STUDY\EDU MATHEMATICS (Yes relative to no)	-10.26***	3.55
	MATH\HOW OFTEN TEACHER GIVE YOU HOMEWORK/MATHEMATICS (More than average)	-0.950	6.12
	MATH\EMPHASIS\NATIONAL OR REGIONAL TESTS (More than average)	1.020	-8.76
	MATH\EMPHASIS\CLASSROOM TESTS (Less than average)	5.22**	-6.92
	MATH\EMPHASIS\ASSESSMENT OF WORK (Less than average)	-2.85*	7.97
Teaching efficacy - Practice - Approach to teaching and learning - Assessment	MATH\HOMEWORK\CORRECT OWN HOMEWORK (Less than average)	-4.81**	-13.98
	MATH\TOPIC\ALGEBRA\FUNCTIONS (Mostly taught the year before, relative to Mostly taught this year)	1.850	13.06
	MATH\TOPIC\ALGEBRA\SIMPLIFYING (Mostly taught the year before, relative to Mostly taught this year)	-5.95***	11.30
	MATH\TOPIC\NUMBER\COMPUTING (Mostly taught the year before, relative to Mostly taught this year)	-1.86*	-13.74
	MATH\TOPIC\DATA\CHARACTERISTICS DATA (Mostly taught the year before, relative to Mostly taught this year)	3.67*	6.03
	MATH\TOPIC\GEOMETRY\CARTESIAN PLANE (Mostly taught the year before, relative to Mostly taught this year)	-5.09**	9.99
	MATH\TOPIC\GEOMETRY\APP MEASUREMENT (Mostly taught the year before, relative to Mostly taught this year)	-2.53*	-8.71
Teaching efficacy - Practice - Approach to teaching and learning - Curriculum	MATH\TOPIC\GEOMETRY\CONGRUENT FIGURES (Mostly taught the year before, relative to Mostly taught this year)	-1.470	-5.05

Theme	Representative question	Full std effect (absolute value)	Partial std effect (absolute value)
	MATH\TOPIC\ALGEBRA\PROPERTIES OF FUNCS (Mostly taught the year before, relative to Mostly taught this year)	-4.28**	-12.06
	MATH\TOPIC\ALGEBRA\NUMERIC (Mostly taught the year before, relative to Mostly taught this year)	3.76*	8.00
	MATH\PERMITTED TO USE CALCULATORS (Yes, restricted relative to Yes, unrestricted)	-3.43*	-7.43
	MATH\ASK STUDENTS\SAME ABILITIES GROUPS (More than average)	-8.34**	5.50
	MATH\ASK STUDENTS\MIXED ABILITIES GROUPS (Less than average)	-2.230	-7.55
	GEN\HOW OFTEN\CLASSROOM DISCUSSIONS (Less than average)	-1.30	-12.56
	GEN\HOW OFTEN\CHALLENGING EXS (More than average)	5.390	9.19
	MATH\ACCESS TO COMPUTER\EACH STD HAS A COMPUTER (Yes relative to no)	-0.590	3.13
	MATH\COMPUTER TABLET ACTIVITIES\DO PROCEDURES (Yes relative to no)	-0.760	4.73
	MATH\HOW OFTEN USE CALC\CHECK ANSWER (More than average)S	1.520	-5.18
	MATH\ASK STUDENTS\WRITTEN TEST OR QUIZ (More than average)S	-0.650	-7.44
	MATH\ASK STUDENTS\NO OBVIOUS SOLUTION (More than average)S	-4.60	-9.82
Teaching efficacy - Practice - Approach to teaching and learning - Instructional approaches	MATH\ASK STUDENTS\WORK IN WHOLE CLASS (Less than average)	-9.99***	4.73
Teaching efficacy - Practice - Approach to teaching and learning - Lesson planning and collaboration	GEN\INTERACTIONS TEACHERS\WORK TOGETHER (More than average)	-8.76***	-9.55
	GEN\HOW FREQUENTLY\MEANING AND PURPOSE (Disagree more than average)	-7.53***	-8.01

Theme	Representative question	Full std effect (absolute value)	Partial std effect (absolute value)
	GEN\THINKING ABT CURR SCH\FEEL SAFE (Disagree more than average)	6.8**	-15.04
	MATH\PREPARED\DATA\INTERPRETING DATA (Less than average)	-1.510	6.11
	MATH\PREPARED\GEOMETRY\GEOMETRIC PROPERTIES (Less than average)	3.81**	-10.12
	MATH\PREPARED\ALGEBRA\PROPERTIES OF FUNCS (Less than average)	2.580	-7.85
	MATH\PREPARED\NUMBER\COMPARE ORDER NUMBERS (Less than average)	-5.37***	-8.82
Teaching efficacy - Practice - Wellbeing and development - Engagement and wellbeing	MATH\CONFIDENT\ASSESS COMPREHENSION (Less than average)	6.160	-5.96
Teaching efficacy - Practice - Wellbeing and development - Professionalism	GEN\DEGREE PROBS TEACH\ARRIVING LATE AT SCHOOL (More than average)	-4.26*	-4.55
	MATH\PROF DEVELOPMENT\MATH CONTENT (Yes relative to no)	-3.840	-3.49
Teaching efficacy - Practice - Wellbeing and development - Support and development	GEN\LANGUAGE DIFFICULTIES (Yes relative to no)	-6.35***	-5.36

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of TIMSS data

Table C.7: Standardised effect sizes of drivers of school quality on TIMSS year 8 science scores

Theme	Representative question	Full std effect	Partial std effect
	GEN\THINKING ABT CURR SCH\STUD BEHAVE (Disagree more than average)	-1.240	-14.15
	GEN\LIMIT TEACHING\MENTAL DISABILITIES (Yes relative to no)	1.61**	9.25
	GEN\LIMIT TEACHING\PHYSICAL DISABILITIES (Yes relative to no)	-0.230	-8.33
Classroom organisation and environment - Environment - Order and cohesion -	GEN\LIMIT TEACHING\LACK OF NUTRITION (More than average)	-2.73***	-14.64
Classroom organisation and environment - Organisation - Compositional strategies of classroom organisation -	GEN\NUMBER OF STUDENTS IN THE CLASS	5.7***	5.29
Classroom organisation and environment - Organisation - Practice driven organisation -	GEN\STUDENT ACHIEVEMENT USED TO ASSIGN\MATH (Yes relative to no)	6.5***	3.48
Classroom organisation and environment - Organisation - Teacher working conditions -	GEN\AGREEMENT\NEED MORE TIME TO ASSIST (Agree less than average)	1.45*	5.25
	GEN\SHORTAGE\MATH\TEACH SPEC MATH (More than average)	-1.870	-12.53
Resourcing - Curriculum and staff based - School based -	GEN\SHORTAGE\SCI\TEACH SPEC SCIEN (More than average)	-1.970	-8.54
	GEN\SEVERITY PROBLEM\CLASSRMS NOT CLEANED (More than average)	-1.36*	-4.50
	GEN\SEVERITY PROBLEM\MATERIAL UNAVAILABLE (More than average)	0.450	-9.25
	GEN\EXISTING SCIENCE LABORATORY (Yes relative to no)	0.60	3.12
Resourcing - Material based - Infrastructure	GEN\SHORTAGE\GEN\HEATING SYSTEMS (More than average)	0.610	-4.13
Resourcing - Material based - Teaching and learning -	GEN\SHORTAGE\MATH\CONCRETE OBJECTS (More than average)	-5.78***	-8.99

Theme	Representative question	Full std effect	Partial std effect
	GEN\BOOKS IN LIBRARY\PRINT (More than average)	-4.270	-6.31
	GEN\SHORTAGE\SCI\LIBRARY RESOURCES (More than average)	5.69**	-7.32
	GEN\SHORTAGE\GEN\INSTRUCTIONAL MATERIAL (More than average)	-2.890	-8.65
	GEN\SEVERITY PROBLEM\INADEQUATE TECH RESOURCES (More than average)	1.360	-5.89
Resourcing - Material based - Technology -	GEN\SHORTAGE\SCI\COMPUTER SOFTWA (More than average)	-1.620	-11.14
	GEN\SCH CHARACTER\PARENTAL COMMITMENT (More than average)	2.69*	9.49
School leadership, governance and culture - Parental culture and involvement - -	GEN\CHARACTERIZE\PARENTAL INVOLVEMENT (More than average)	2.02*	7.70
School leadership, governance and culture - Policy, mission and goals - Development and collaboration -	GEN\CHARACTERIZE\COLLABORATION TO PLAN (More than average)	2.25**	-3.18
	GEN\USE INCENTIVES\MATH (Yes relative to no)	-8.78***	-3.58
School leadership, governance and culture - Policy, mission and goals - Hiring and recruitment -	GEN\FILL TEACHING VACANCIES\SCIENCE (More difficulty than average)	2.83*	-6.66
	GEN\THINKING ABT CURR SCH\RULES ENFORCEMENT (Disagree more than average)	-4.29***	-10.83
	GEN\FREE MEALS\BREAKFAST (Yes relative to no)	0.480	-13.73
	GEN\ASSIST WITH SCHOOLWORK (Yes relative to no)	-4.89***	-5.83
School leadership, governance and culture - Policy, mission and goals - School level approach to practice - -	GEN\THINKING ABT CURR SCH\CLEAR RULES (Disagree more than average)	1.10	-7.76
School leadership, governance and culture - Policy, mission and goals - Transparency and clarity -	GEN\CHARACTERIZE\CLARITY OF OBJECTIVES (Lower than average)	1.55*	-6.89
	GEN\YEARS PRINCIPAL ALTOGETHER	2.95**	4.14

Theme	Representative question	Full std effect	Partial std effect
School leadership, governance and culture - Principal attributes, culture and integration - Attributes	GEN\HIGHEST LEVEL OF FORMAL EDUCATION (Higher than average)	0.28***	-10.78
School leadership, governance and culture - Principal attributes, culture and integration - Role -	GEN\CHARACTERIZE\SUPPORT FOP PROF DEVELOPM (Higher than average)	1.40	4.01
	GEN\CHARACTERIZE\TCHS DEGREE OF SUCCESS (Higher than average)	1.37*	-6.97
School leadership, governance and culture - Staff culture and involvement - -	GEN\SCH CHARACTER\TCH UNDERSTANDING (Lower than average)	1.540	-8.40
Teaching efficacy - Attributes - Characteristics - Personal characteristics	GEN\SEX OF TEACHER (Male relative to female)	1.52*	2.85
	GEN\YEARS BEEN TEACHING	0.70	2.29
Teaching efficacy - Attributes - Qualifications - Level of teacher education	GEN\LEVEL OF FORMAL EDUCATION COMPLETED (Higher than average)	-1.07**	4.15
Teaching efficacy - Attributes - Qualifications - Type of teacher education	GEN\MAJOR AREA OF STUDY\BIOLOGY (Yes relative to no)	3.12**	1.28
	GEN\MAJOR AREA OF STUDY\MATHEMATICS (Yes relative to no)	-0.670	-2.21
Teaching efficacy - Practice - Approach to teaching and learning - Assessment	SCI\HOW OFTEN TEACHER GIVE YOU HOMEWORK/SCIENCE (More than average)	-2.920	-7.14
	SCI\EMPHASIS\CLASSROOM TESTS (Less than average)	0.780	-5.97
	SCI\HOMEWORK\USE FOR GRADES (Yes relative to no)	-0.290	-5.09
Teaching efficacy - Practice - Approach to teaching and learning - Curriculum	SCI\TOPIC\EARTH\EARTHS RESOURCES (Mostly taught the year before relative to Mostly taught this year)	0.670	4.01
	SCI\TOPIC\PHY\ENERGY FORMS (Mostly taught the year before relative to Mostly taught this year)	0.740	5.41
	SCI\TOPIC\PHY\PHYSICAL STATES (Mostly taught the year before relative to Mostly taught this year)	1.240	6.09

Theme	Representative question	Full std effect	Partial std effect
	SCI\TOPIC\CHEM\MIXTURES AND SOLUTIONS (Mostly taught the year before relative to Mostly taught this year)	-1.43*	5.74
	SCI\TOPIC\CHEM\PHY CHEM PROPERTY MATTER (Mostly taught the year before relative to Mostly taught this year)	-0.82*	-6.25
	SCI\TOPIC\CHEM\CLASSIFICATION (Mostly taught the year before relative to Mostly taught this year)	-0.060	7.19
	SCI\TOPIC\BIO\DIFFERENCES TAXONOMIC GROUPS (Mostly taught the year before relative to Mostly taught this year)	-0.670	-7.39
	SCI\TIME SPENT SCIENCE	4.56***	3.50
	SCI\TOPIC\EARTH\EARTH IN SOLAR SYSTEM (Mostly taught the year before relative to Mostly taught this year)	-1.220	-5.04
	SCI\TOPIC\EARTH\EARTHS PROCESSES (Mostly taught the year before relative to Mostly taught this year)	1.340	5.23
	SCI\TOPIC\EARTH\EARTHS STRUCTURE (Mostly taught the year before relative to Mostly taught this year)	-1.82**	4.18
	SCI\TOPIC\PHY\FORCES AND MOTION (Mostly taught the year before relative to Mostly taught this year)	-2.820	-7.18
	SCI\TOPIC\PHY\ELECTRIC CIRCUITS (Mostly taught the year before relative to Mostly taught this year)	1.46**	5.80
	SCI\TOPIC\PHY\BASIC PROPERTIES OF LIGHT (Mostly taught the year before relative to Mostly taught this year)	-0.420	5.91
Teaching efficacy - Practice - Approach to teaching and learning - Instructional approaches	SCI\COMPUTER TABLET AVAILABILITY DURING SCI (Yes relative to no)	-0.870	3.10
Teaching efficacy - Practice - Approach to teaching and learning - Lesson planning and collaboration	GEN\INTERACTIONS TEACHERS\COLLABORATE (More than average)	-1.510	-4.20

Theme	Representative question	Full std effect	Partial std effect
	SCI\CONFIDENT\ENGAGE STUDENTS INTEREST (More than average)	0.780	-6.61
	GEN\HOW FREQUENTLY\SATISFIED TEACHER (More than average)	0.80	-7.70
	SCI\PREPARED\EARTH\EARTHS PROCESSES (Less than average)	-1.1*	-5.50
	SCI\PREPARED\CHEM\CHEMICAL CHANGE (Less than average)	1.41*	-6.64
Teaching efficacy - Practice - Wellbeing and development - Engagement and wellbeing	SCI\PREPARED\BIO\CELLS AND THEIR FUNCTION (Less than average)	-0.40	5.47
Teaching efficacy - Practice - Wellbeing and development - Support and development	SCI\PROF DEVELOPMENT\SCI ASSESSMENT (Yes relative to no)	-3.97***	-1.84
	SCI<PROF DEVELOPMENT> HOURS	1.760	4.36

Significance level: *** 1 percent; ** 5 percent; * 10 percent.

Source: Deloitte Access Economics analysis of PISA data

C.4. Geographic analysis

Table C.8: Relative importance of school quality themes by geography – PISA (maths)

Themes	Metro	Provincial	Remote
Accountability	0.01%	0.10%	0.62%
Autonomy	0.32%	1.04%	1.53%
Classroom organisation and environment - Environment	4.77%	6.34%	7.84%
Classroom organisation and environment - Organisation	0.18%	0.35%	1.53%
Resourcing - Material based	0.26%	0.60%	1.53%
Resourcing - Curriculum and staff based	n/a	n/a	n/a
School leadership, governance and culture*	2.09%	4.66%	1.53%
Teaching efficacy - Attributes	0.06%	0.11%	0.43%
Teaching efficacy - Practice	5.89%	7.69%	16.14%
All school drivers	13.53%	21.74%	24.42%

Source: Deloitte Access Economics analysis of PISA data

Table C.9: Relative importance of school quality themes by geography – TIMSS (maths, year 8)

Themes	Metro	Provincial	Remote
Accountability	n/a	n/a	n/a
Autonomy	n/a	n/a	n/a
Classroom organisation and environment - Environment	3.65%	3.25%	0.30%
Classroom organisation and environment - Organisation	3.59%	4.07%	1.46%
Resourcing - Material based	2.60%	7.31%	n/a
Resourcing - Curriculum and staff based	0.49%	0.02%	n/a
School leadership, governance and culture*	4.86%	9.08%	1.46%
Teaching efficacy - Attributes	0.80%	3.94%	1.46%
Teaching efficacy - Practice	15.00%	13.55%	12.63%
All school drivers	30.30%	27.02%	14.09%

Source: Deloitte Access Economics analysis of TIMSS data

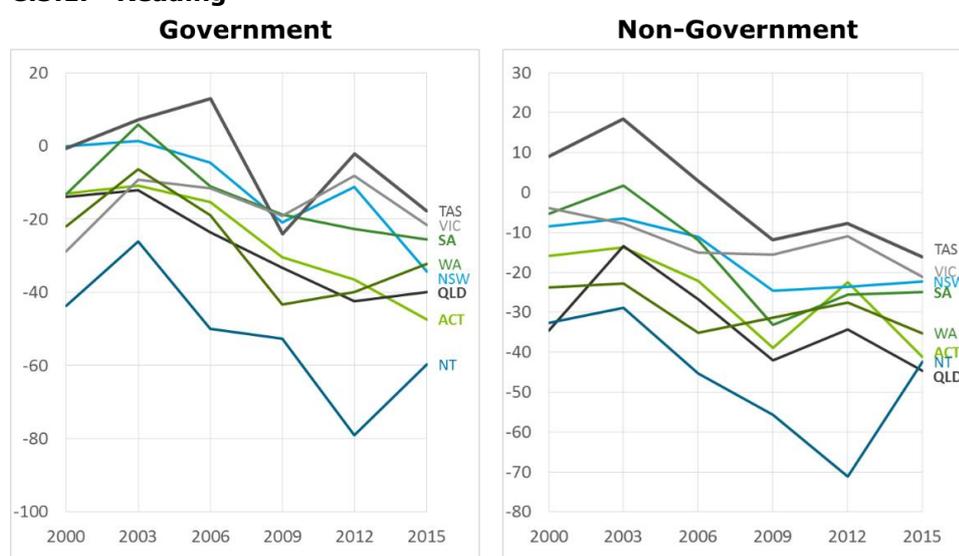
Table C.10: Relative importance of school quality themes by geography – TIMSS (maths, year 4)

Themes	Metro	Provincial	Remote
Autonomy	n/a	n/a	n/a
Classroom organisation and environment - Environment	0.92%	0.61%	n/a
Classroom organisation and environment - Organisation	0.94%	1.21%	n/a
Resourcing - Material based	1.27%	0.85%	n/a
Resourcing - Curriculum and staff based	0.60%	0.59%	n/a
School leadership, governance and culture*	2.66%	1.33%	0.48%
Teaching efficacy - Attributes	0.38%	0.06%	0.47%
Teaching efficacy - Practice	4.68%	3.08%	0.48%
All school drivers	10.83%	6.81%	0.96%

Source: Deloitte Access Economics analysis of TIMSS data

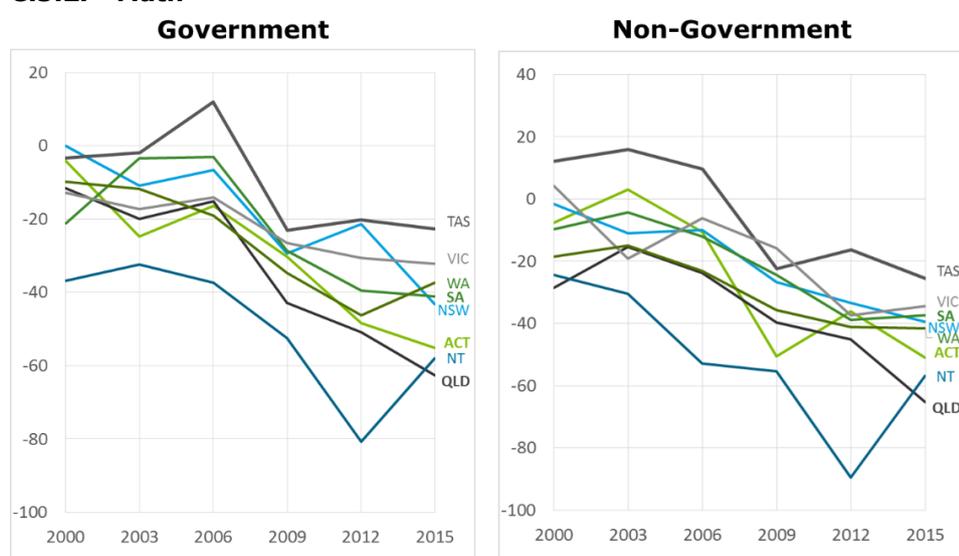
C.5. PISA, Relative value add by test domain, jurisdiction and school system over time with student and school characteristic controls, 2000-2015

C.5.1. Reading



Source: Deloitte Access Economics analysis of PISA data

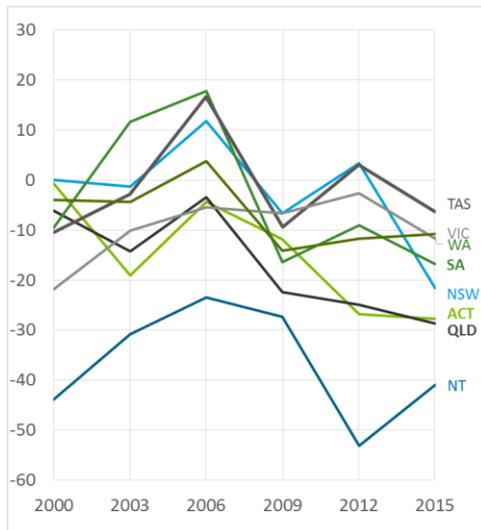
C.5.2. Math



Source: Deloitte Access Economics analysis of PISA data

C.5.3. Science

Government



Non-Government



Source: Deloitte Access Economics analysis of PISA data

Appendix D: Providing a link between schooling system settings and performance across Australian jurisdictions

This section outlines the findings from the performance analysis of Australian jurisdictions on PISA and TIMSS, and seeks to link the variations in this performance—across jurisdictions, and over time—to observed historical differences in the policies and practices of schooling systems across Australian jurisdictions.

As distinct from the analysis above, which considers some aspects of system policy in a ‘bottom-up’ fashion, this section seeks to consider how differences in system settings and initiatives are related to system level performance in a ‘top-down’ manner. This is achieved by looking at the performance of Australia’s schooling jurisdictions over time in the context of their system settings and policies, drawing on desktop evidence from outside of the TIMSS and PISA datasets.

D.1. Trends in performance across Australian jurisdictions

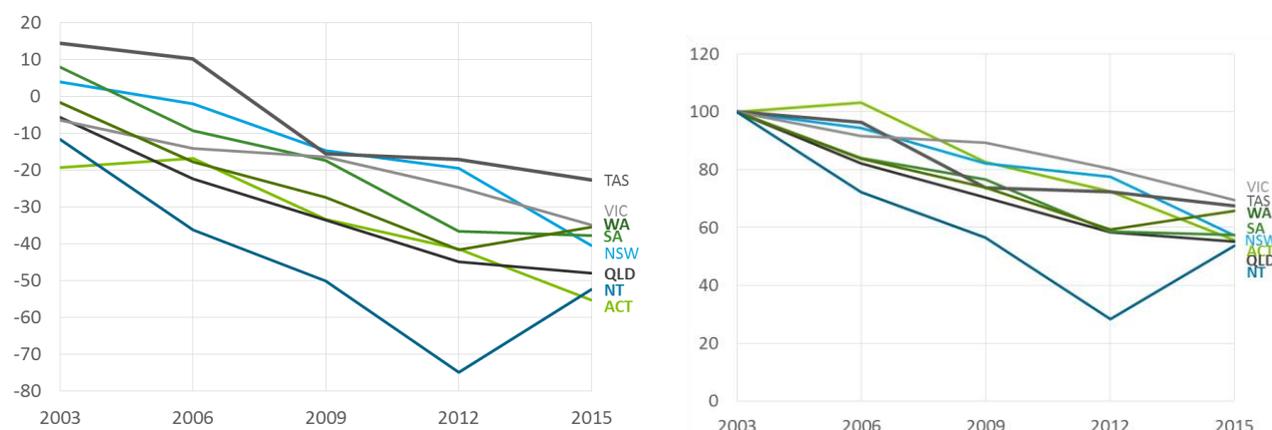
Chart D.1 below shows the relative value-add of government schools in each jurisdiction on PISA scores since 2003, compared to the performance of NSW non-government schools in 2003.³⁹ After accounting for the inherent characteristics of students and schools (through regression modelling of the impact of these factors on student outcomes) it is possible to isolate the extent to which the performance of each jurisdiction differs in a way that materially affects student outcomes. These relative effects on student outcomes may be associated with system-level differences between schooling systems.

However, it should be noted that there are other factors that may explain these differences in measured performance, including contextual factors that cannot be fully accounted for in the regression analysis. Importantly, this analysis should not be used to rank the relative effectiveness of different jurisdictions’ schooling system. Rather, the analysis has been developed purely for the purpose of supporting exploratory research into the possible impact of policy interventions over recent years. In addition to the analysis which compares overall value-added of government schools across jurisdictions, Chart D.2 shows the relative change in value add from each jurisdiction’s own score in 2003.

More detailed charts showing the relative value add of each jurisdiction in each PISA domain are provided in Appendix C.

³⁹ It should be noted that NSW non-government schools are simply used as a reference category and this analysis is not intended and cannot be used to make assessments of the relative performance of the government and non-government schooling systems in Australia.

Chart D.1: Trends in PISA value-add by state, 2003-2015 Chart D.2: Trends in PISA value-add by state, 2003-2015 (score in 2003 = 100)



Source: Deloitte Access Economics analysis of PISA data

Key observations that can be made from this evidence include:

- While controlling for contextual factors, the **differences in performance between jurisdictions are smaller** than they appear otherwise (that is, as outlined in Chart 3.3 Chart 3.4 earlier in this report, where student and school context is not accounted for). This suggests that the difference in the performance between jurisdictions is not as material as un-adjusted average student test scores would suggest
- **Victoria** has had the **smallest decline** across the time period across subject areas.
- **Tasmania** has performed **at or above the level of other jurisdictions** over the time period. It has also tended to have the greatest improvement, or smallest decline, in value-add among the states.
- The **Northern Territory improved dramatically** between 2012 and 2015, while **Western Australia** also recorded a **significant improvement** in that time frame (noting that the small sample size in the Northern Territory limits the precision of this analysis).
- Over the entire time period, other states and territories declined by roughly the same amount relative to their original position.

Limitations of value-add as a longitudinal performance measure

Value-add is only one way of thinking about school outcomes and performance. This measure has the benefit of controlling for various school- and student-level contextual factors (namely, those outlined in section 4.1 above). However, it also has limitations as a measure of performance over time. In particular, neither TIMSS nor PISA use the same sample of students in each test – they test a different cohort of students in the same year levels (for TIMSS) or age (in PISA) from different schools. Accordingly, results are highly sensitive to the approach taken to sampling. As a result, value-add provides no more than an indication of the average levels and trends in performance between jurisdictions. It should be examined in the context of outcome measures from other sources.

In the following section, significant differences between these jurisdictions' schooling systems and over time are outlined, which may contribute to these trends. The aspects of difference are included in this analysis are selected based on a process of identification and prioritisation, as set out in section 3.4.2 of the methodology chapter of this report.

D.2. Identification of key system settings

Deloitte Access Economics has identified several historical differences between schooling systems, supported by preliminary research from the Department in line with the methodology set out in section 3.4.2 of this report. Permanent and timing differences have been prioritised by the:

- Quality of evidence;
- Timing of expected effects (that is, whether they are expected to occur within the period of available evidence); and
- Strength of expected impact (based on the literature).

To aid in the understanding of the different forms which government initiatives may take, Figure D.1 below provides a framework through which governments affect school practice and, ultimately, student outcomes. In this framework:

- Governments have direct power over **system settings** – the framework in which schools operate, the goals which they pursue, and the processes for demonstrating that schools are achieving those goals.
- Governments also establish **enabling initiatives** that provide guidance to schools in pursuing system-level goals effectively.
- Schools undertake **school interventions** to change their school practice to pursue system-wide goals. These include measures like teacher professional learning and pedagogy. However, they do not undertake these activities in isolation. They select interventions from those available from government authorities and those available on the market. Importantly, these interventions generally happen at a school level, rather than at a system-wide level. Schools are most aware of the particular needs of their cohort and few interventions are appropriate for implementation across the entire schooling system.
- On a day-to-day basis, schools are responsible for pursuing these goals. How effectively they achieve these goals is determined through their overall **school practice**: their classroom practice, school culture and teaching efficacy.

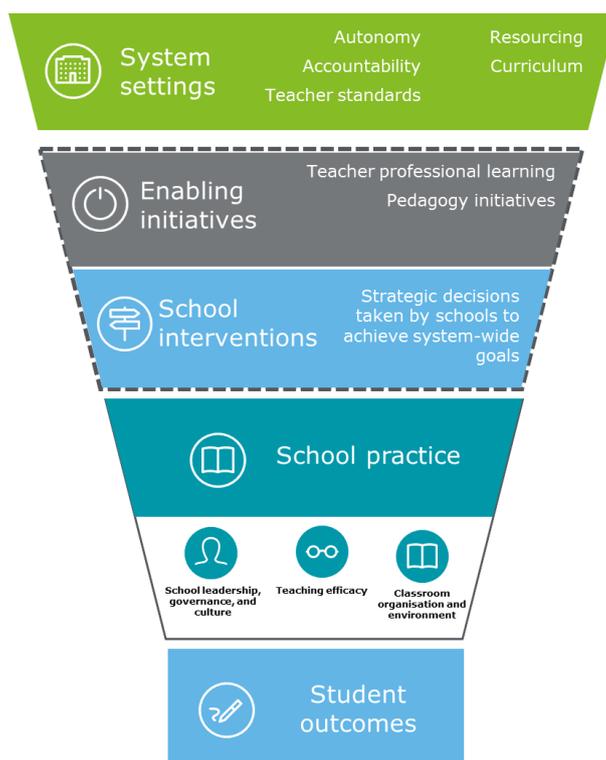
In this framework, governments have two tools to manipulate in achieving school outcomes: (1) system settings; and (2) enabling initiatives. This section examines each of these in turn.

Historical analysis of system settings

The following section is primarily an exploratory exercise. Having identified an evidence base for particular practices in the earlier literature review, the section maps the historical differences between jurisdictions. This is done with the aim of providing some understanding of why variation exists in outcomes across jurisdictions. It does not aim to provide conclusive evidence that certain practices are or are not effective in the Australian context – such conclusions could only be drawn from controlled trials or other evaluations.

The policies and initiatives discussed are not necessarily current or best practice. Rather, they serve as illustrations of historic practice. Changes at a system level may take several years to have an impact on student outcomes. Accordingly, this research has been driven by past reforms. In some cases where this historical mapping exercise does not reflect current practice, recent reforms are noted.

Figure D.1: Framework relating system settings and initiatives to aspects of school practice



Source: Deloitte Access Economics

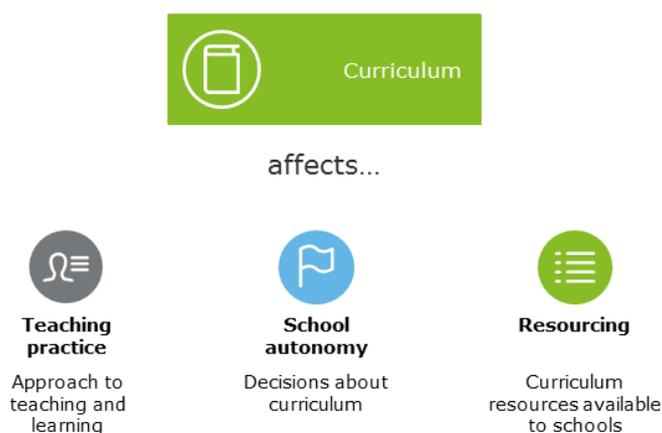
D.3. System settings

Curriculum

Curriculum—the content of the education which teachers deliver in schools—forms a fundamental part of the schooling system. Accordingly, it can have an impact on student outcomes across a number of domains (see Figure D.2). In particular, of the domains identified in this study’s framework of themes of school quality outlined in section 2.2 above, curriculum can influence aspects of:

- **Teaching practice** – by setting areas of focus for the classroom;
- **School autonomy** – by determining the extent to which schools can structure their teaching time; and
- **Resourcing** – by providing the building blocks of lessons made available to schools.

Figure D.2: Impact of curriculum on school quality



Source: Deloitte Access Economics

In 2008, the Australian Government established the Australian Curriculum, Assessment and Reporting Authority (ACARA) which began the task of creating a national curriculum for Australian schools in English, mathematics, science and history for students between Foundation and Year 10. The Review of the Australian Curriculum (Australian Government, 2014) noted that this push in Australia came alongside a trend of curriculum review in many parts of the world. The Review noted observations of the OECD that the highest-performing curricula:

- Develop the *conceptualisation capacity* of students;
- Contain *prescriptive content* but leave pedagogy, assessment and reporting to teachers and schools (in varying degrees);
- Maintain a *strong curriculum framework and accountability* in an environment of school autonomy;
- Teach *not just knowledge, but competencies*;
- Rely on a *mixture of student-centred learning, explicit teaching and project-based activity*; and
- Contain both *formative and summative* 'high-stakes' assessment.

These are all objects of the Australian Curriculum to varying degrees—in particular, the curriculum has a focus on the development of general skills and abilities, rather than focusing only on discipline-based skills.

The first four learning areas of the Australian Curriculum—English, Mathematics, Science and History—were endorsed in December 2011 for students between Foundation and Year 10. However, the Curriculum was not implemented at once, but rather at different points in time between:

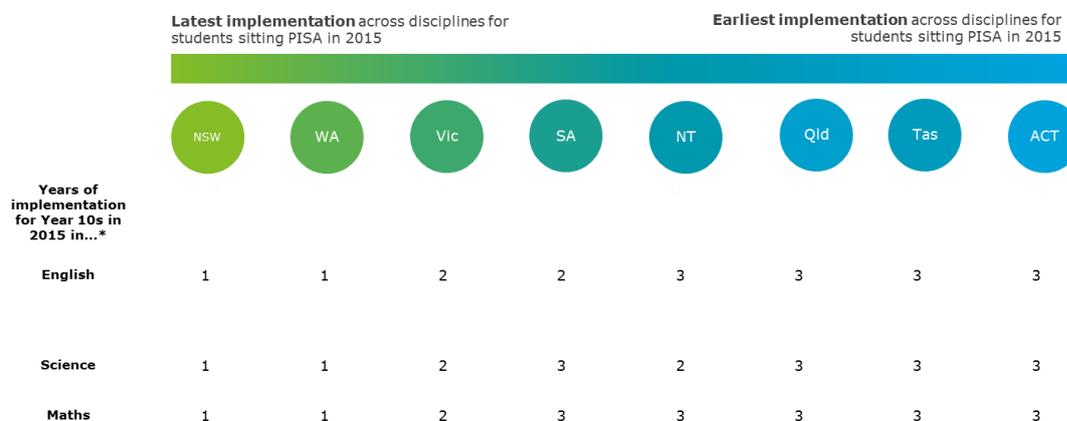
- Learning areas;
- States and Territories;
- School year levels; and
- School sectors.

Because of implementation across all Australian schools, and the ability to track the sequence of implementation over time, it is possible to concord the implementation of the Australian curriculum against performance on PISA and TIMSS. If the Australian Curriculum has had a positive effect in developing the capacity of students to tackle the types of problems in the two tests, it would be expected that this difference only appear in states which had implemented the Australian Curriculum *before* the test dates observed in the PISA and TIMSS datasets.

Using published implementation dates from ACARA and state and territory education departments, it is possible to identify the point in time at which the Australian Curriculum had been implemented in each of its core learning areas. Evidence from the literature review also suggests that the effect of improved curriculum is *cumulative*—that is, that the gain from improved curriculums increases

through several years of schoolings. If the Australian Curriculum has had an impact on student outcomes, a more significant improvement—or smaller decline—would be expected in states where it was implemented than those where it was not.

Figure D.3: Mapping the implementation of the Australian Curriculum across Australian jurisdictions



*Measures the number of years prior to 2015 completed under the Australian Curriculum by students completing Year 10 in 2015 in government schools.

Source: ACARA (2014), NSW Board of Studies (2012)

Noting the trends in PISA results discussed in Section 4.3.1 above, there is no clear sign that jurisdictions that implemented the Curriculum earlier performed better in PISA. The ACT, Tasmania, and Queensland were among the earliest jurisdictions to implement the curriculum across a number of domains, yet all declined in the value-added measure (by varying degrees) between 2012 and 2015 after controlling for contextual factors.

This is not to say that implementing the Australian Curriculum has not had a beneficial effect on student outcomes in Australia. There are a number of caveats that prevent drawing strong conclusions from any comparisons of states based on implementation dates:

- The implementation of the Curriculum was targeted as much, if not more, at making Australian school curriculums uniform as it was at deploying a high-quality curriculum. To expect a clear impact from its implementation presupposes that each Australian jurisdiction's existing curriculum was not already teaching the PISA domains at international standards.
- Apart from the ACT, all Australian jurisdictions made modifications to the Australian Curriculum to suit local needs. These jurisdiction-level changes cannot be separated from the benefit of the national push for a uniform curriculum.
- Implementation of the Curriculum may have occurred significantly later than published implementation dates.
- Further, primary responsibility for the implementation of the Curriculum fell to the hands of schools, which may vary significantly in their practices.

Autonomy

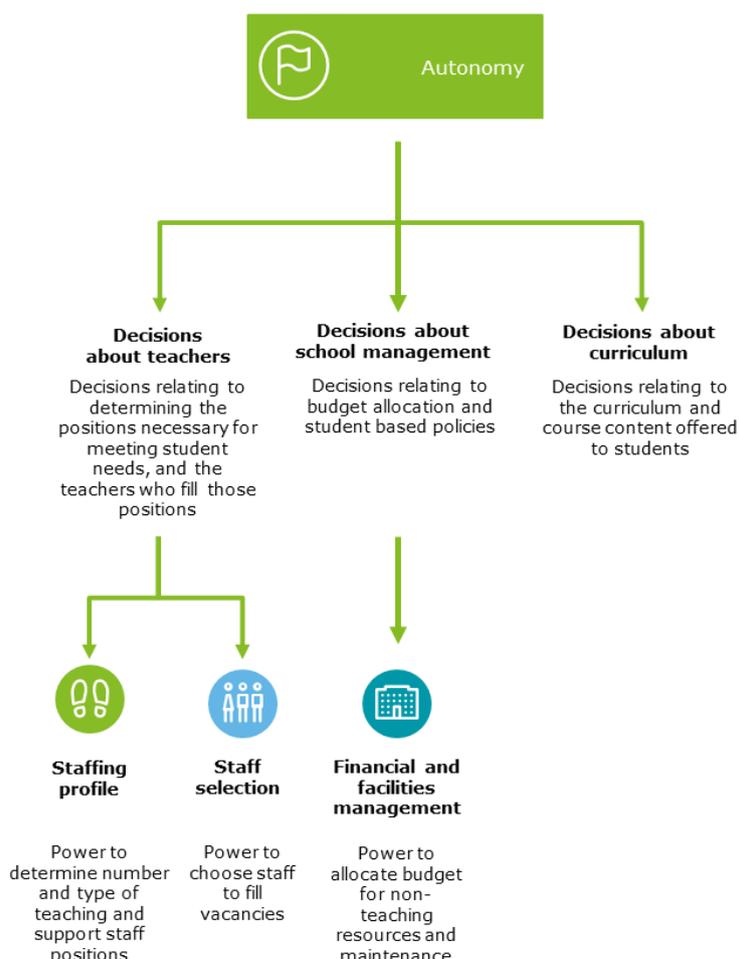
School autonomy—that is, greater local decision-making at the level of the school and school community—has been identified in the literature as a crucial factor to good student outcomes (Gonski, 2011). Autonomy over process and personnel decisions has been found to encourage higher performance at a school level. However, effective accountability mechanisms at a system-level are necessary to realise the potential of school autonomy, to ensure that schools are making decisions that make progress towards system goals (Hanushek and Woessmann, 2014; Hanushek, Link and Woessmann, 2013).

While the empirical analysis outlined in section 4.3 above considered measures of system autonomy identified by schools in PISA and TIMSS, here consider measures of autonomy across Australian jurisdictions are considered through a consideration of the differing nature of system

settings drawn from desktop research and evidence compiled from sources outside of PISA and TIMSS.

Autonomy is not a one-size-fits-all reform. Many complex decisions are made on a day-to-day basis throughout schooling systems, and accordingly, there are different types of autonomy that may be provided to schools (illustrated in Figure D.4 below).

Figure D.4: Types of school autonomy



Source: Deloitte Access Economics

It is also important to distinguish between **structural** and **professional autonomy** in schools. Schools may be given *structural authority* to make particular decisions about their operations and their classroom practice. However, in order to see improvements in results, they need the *professional capacity* to exercise that authority in pursuit of key system goals.

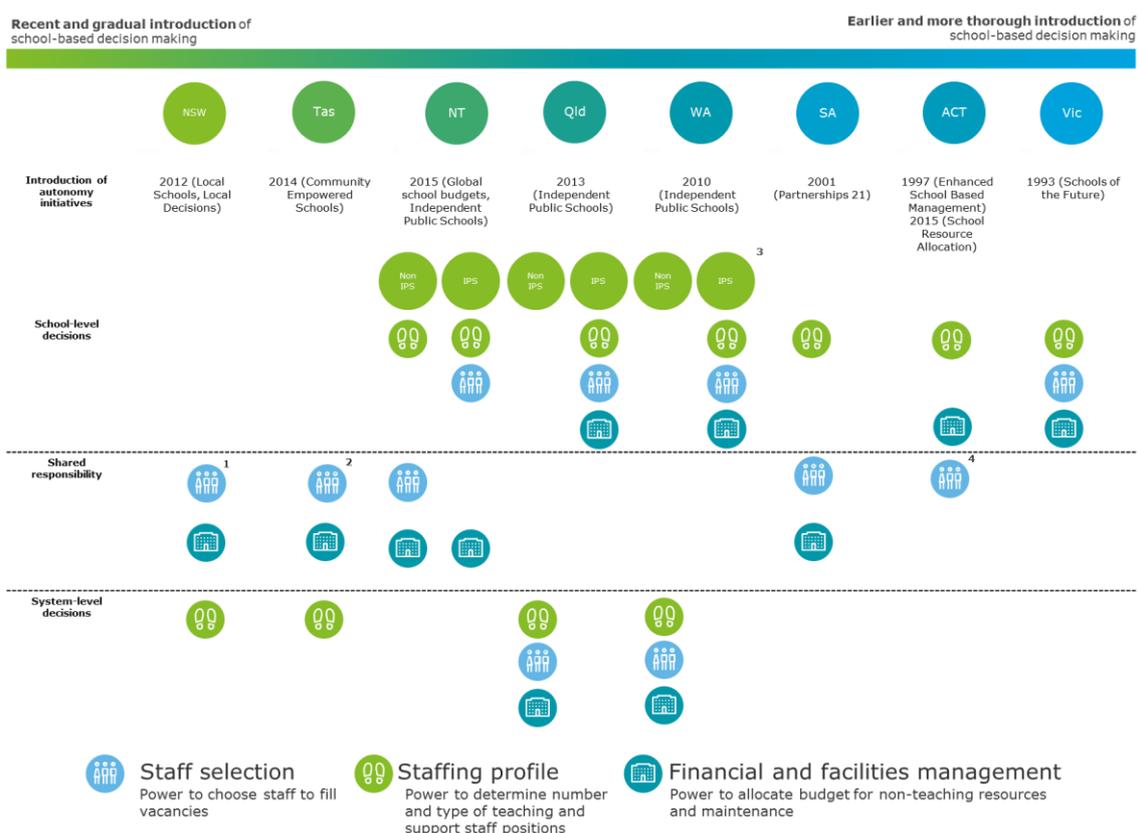
Australian school systems have had varying degrees of autonomy for some time. There has been no attempt to comprehensively catalogue the relative autonomy of government or non-government schools, though there have been a number of separate studies which make some effort to consider the relative autonomy of each state's government schooling systems. Although devolution of decision-making power to government schools was first seriously floated in 1973 in the Karmel Report, limited progress was made until the 1990s.⁴⁰ Each Australian jurisdiction can be grouped into one of three categories as to their relative focus on school autonomy:

⁴⁰ Interim Committee for the Australian Schools Commission (1973), cited in Caldwell et al (2015).

- **Victoria** has offered a unique environment for schools in terms of autonomy since the early 1990s. More than 90 per cent of the state’s recurring annual budget is decentralised for schools to determine their own approach to staffing and administration. **South Australia** and the **ACT** also devolved powers to a greater extent than other jurisdictions in the early 2000s, though central authorities still have a greater say over school staffing decisions and management than in Victoria.
- **Western Australia**, in 2010, **Queensland**, in 2013, and the **Northern Territory**, in 2015, implemented a system of Independent Public Schools. Under this system certain schools with demonstrated capacity to make the most of autonomy are granted control over their staffing profile and selection processes. Other schools are generally centrally staffed—in part because of the difficulty of attracting staff to regional and rural schools in these jurisdictions. In Western Australia, by 2015, around 70 per cent of public school students and teachers were in Independent Public Schools
- The remaining jurisdictions—**Tasmania** and **New South Wales**—have only recently implemented autonomy initiatives. Under the Commonwealth Independent Public Schools National Partnerships (Australian Government, 2017), these states are implementing further initiatives to empower schools at a local level. These initiatives are separate to the Independent Public Schools operating in WA, Queensland and the NT.

Figure D.5 maps out the relative levels of autonomy across Australian jurisdictions, across the range of dimensions discussed above.

Figure D.5: Mapping autonomy across Australian jurisdictions



¹ One in two vacancies, after mandatory redeployments such as transfers, are selected by schools.

² School-level staffing decisions must be ultimately approved by the Department of Education.

³ In 2013, Western Australia began implementing several measures to improve school autonomy in non-IPS schools. Schools now operate on a one-line budget and can determine their own staffing profile. This diagram illustrates the position for non-IPS schools prior to 2013.

⁴ Although schools have a general prerogative to select their own staff, the Education Directorate has several powers to place teachers at a system level, including through transfers and at the Director-General’s discretion.

Sources: ACT Directorate of Education and Training (2004), ACT Directorate of Education and Training (2012), Australian Government (2014a), Australian Government (2014b), Australian Government (2017), Caldwell (2015), Kilvert (2001), NSW Department of Education and Communities (2014), Queensland Department of Education and Training (n.d.), South Australian Department for Education and Child Development (2015), University of Melbourne Graduate School of Education (2013), Western Australian Auditor-General (2011)

Victoria has consistently been among the top performers in PISA, and has had a long-term focus on autonomy as a jurisdiction. Western Australia also had a notable improvement between 2012 and 2015, as the proportion of that state’s Independent Public Schools increased. However, the ACT performs poorly in terms of value-add, despite also having an earlier focus on autonomy as Victoria did; similarly, even though the Northern Territory had not made substantial autonomy-focused changes until 2015, it had an even more dramatic improvement than Western Australia.

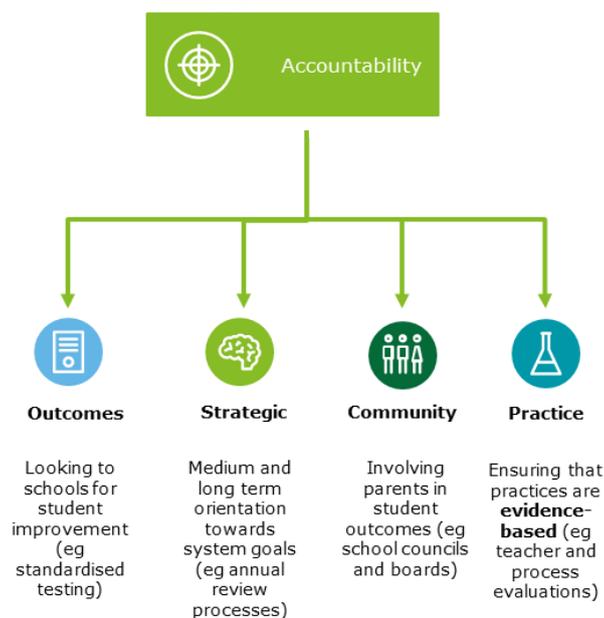
This aligns with earlier findings from PISA and TIMSS (see section 4.2.1) suggesting that autonomy in management and teacher selection has a small correlation with outcomes. In any case, autonomy enables schools to meet local needs more effectively—improving teaching practice by allowing schools to make decisions appropriate for their context. In this way, the impact of autonomy may not appear as a direct result of autonomy, but rather through the indirect avenue of improved teaching practice.

Accountability

Autonomy appears to be an important driver of school quality from the literature. However, if not paired with effective accountability measures, school leaders may not face clear incentives to guide their decision-making. School leaders have the opportunity to develop a culture of accountability within their school by focusing on student improvement at a school level. There are several mechanisms through which schooling systems can render schools accountable for outcomes improvement:

- **Outcomes accountability** focuses on student results and improvement;
- **Strategic accountability** focuses on ensuring that schools have medium- and long-term plans to focus on improvement;
- **Community accountability** focuses on allowing parents and other members of school communities to provide feedback and direction to schools; and
- **Practice accountability** focuses on ensuring that the practices in schools are evidence-based.

Figure D.6: Accountability framework



Source: Deloitte Access Economics

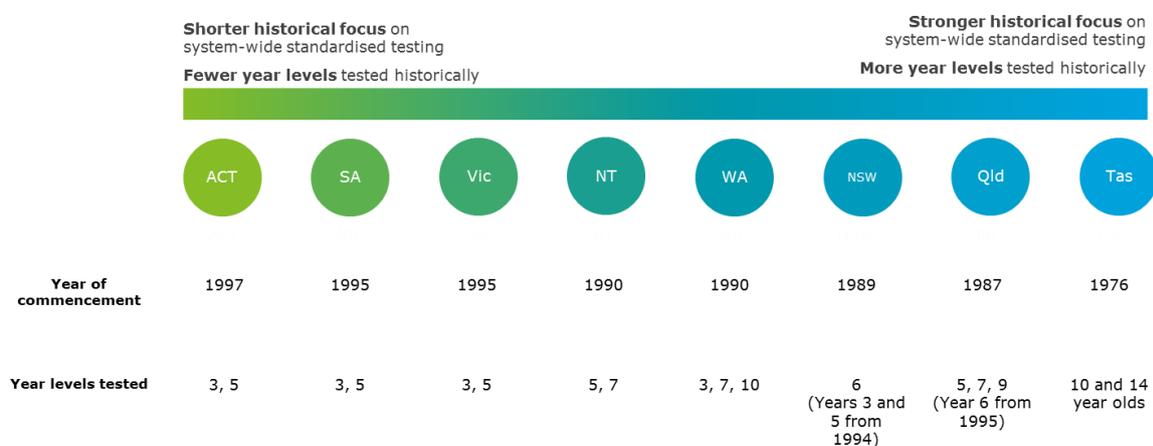
In these ways, accountability *indirectly* drives student outcomes by setting the agenda against which schools direct their resources and classroom practice. Although governments have a role to play in each of these forms of accountability, their goal is most clear in outcomes, strategic and practice accountability. Below, the analysis focus on standardised testing initiatives, the well-

embedded form of outcomes accountability in the Australian schooling system. Standardised testing, as a form of outcomes accountability, drives a school-level focus on student outcomes in several ways:

- It allows for **longitudinal comparisons** at an individual level. Improvements, or declines, in student outcomes between years can be recognised, aiding targeted teaching practices.
- It allows for comparison **between subgroups**, particularly those identified as being in need of additional support – for instance, low socio-economic status students.
- Consistent standardised testing throughout primary and high school encourages **early diagnosis of literacy and numeracy issues**.

In Australia, standardised testing rose to prominence upon the implementation of the National Assessment Plan for Literacy and Numeracy (NAPLAN) in 2008. The aim was to place all students on a single national scale for literacy and numeracy achievement (Fachinetti, 2015). Prior to NAPLAN, each jurisdiction, to varying degrees, had standardised testing initiatives to compare student performance consistently between schools. These initiatives are outlined in Figure D.7 below.

Figure D.7: Mapping standardised testing initiatives



Source: Deloitte Access Economics desktop research

In relation to PISA results over time, it is worth noting that Tasmania, the state with the longest-standing standardised testing initiative, has consistently performed at or above the average level of other jurisdictions. This may not be surprising, given that PISA itself is a standardised test. Nonetheless, it suggests that the system, and its schools and teachers, have had a focus on student outcomes and improvement over a longer time frame than other Australian jurisdictions.

Separate from the historical introduction of standardised testing, it is also worth noting the recent introduction of the Online Literacy and Numeracy Assessment (OLNA) in Western Australia, which has been credited with driving outcome improvement in that state.

Western Australia – Online Literacy and Numeracy Assessment

From 2014, the Western Australian government introduced a minimum literacy and numeracy standard. Students must demonstrate a level of reading, writing and numeracy needed to “meet the demands of everyday life” (Western Australian Government, 2013). Students may demonstrate this standard in one of two ways:

- Students who achieve Band 8 or higher in the reading, writing or numeracy component of their Year 9 NAPLAN test do not need to complete the corresponding section of the OLNA.
- Other students have a total of six opportunities opportunity to sit an Online Literacy and Numeracy Assessment (OLNA): two times per year from Year 10 until Year 12. They need only reach the standard once in each component.

Students must reach this minimum standard by the end of Year 12 to be eligible to receive a high school certificate. Students who do not pass the OLNA instead do receive an official statement of their secondary school achievements to provide to employers or training providers. The Department of Education offers assistance and adjustments to the OLNA for students with a disability.

The OLNA drives improvements in literacy and numeracy in two key ways:

- Most directly, it creates a form of external accountability to ensure that schools are providing support to assist students in reaching the minimum standard by the time they complete Year 12.
- More indirectly, it provides an important incentive for students and schools to improve performance on Year 9 NAPLAN tests, and to use those tests to diagnose any existing literacy and numeracy issues to provide appropriate support to students to assist them in meeting the standard before they complete high school.

28.1% of 2016's Year 12 graduates had achieved the minimum standard in their Year 9 NAPLAN test; more than 90 per cent of those students achieved the minimum standard by Year 12 (Western Australian Government, 2017). The OLNA has been credited with raising NAPLAN performance, with Western Australia recording results above the mean of other states in reading, writing and numeracy (AAP, 2016). The New South Wales Government intends to introduce a similar requirement from 2020 onwards (NSW Education Standards Authority, 2017). This apparent impact is also evident in the 2016 PISA results in reading and numeracy, where Western Australia appears to have had significant recent improvements (see above).

With respect to accountability, there does not appear to be any level of consistency among other jurisdictions, which have on the whole performed or declined at similar levels since 2003. Further, the introduction of NAPLAN testing in 2008 does not seem to have produced a consistent lift in student outcomes. It should be noted that NAPLAN testing cannot be isolated from the number of other changes happening at a jurisdictional level at that time. Nonetheless, accountability serves more as a structure in which schools seek system level gains in outcomes, rather than as a panacea for school improvement.

Industrial arrangements

Industrial arrangements for teachers – their pay and conditions – comprise one of the most significant parts of resourcing provided by schooling systems to schools (in dollar terms). Teachers—as a human resource—can be distinguished from material and curriculum resources in the measurement framework developed in this study.

The direct impact of industrial arrangements is on teacher *attributes*, rather than teaching practice. The pay, conditions and status of teachers relative to other professions have potential to “crowd out” good graduates from the teaching profession (Ingvarson and Rowe, 2007); more competitive salaries help to attract higher-performing high school graduates to the teaching profession. Indeed, for teachers, while their salary on graduation might be competitive with their peers in other professions, salaries in other industries often grow more rapidly with experience (Chevalier and McIntosh, 2001). However, this view also presupposes that teachers' capacity is, to some degree, fixed by their performance in high school, and that salaries dominate among the many and varied incentives for individuals to pursue teaching careers. It is difficult to draw firm conclusions from the literature relating to the impact of teacher industrial arrangements on classroom quality. Nonetheless, industrial arrangements have potential to shape a number of incentives for potential and serving teachers (see Figure D.8 below).

Figure D.8: Industrial arrangements framework for affecting school quality



Source: Deloitte Access Economics

Australia has an older teaching service: the typical Australian teacher has taught for 17 years (OECD, 2013). Changing the incentives for individuals to start a career in teaching will not change the inherent attributes nor the practices of the majority of teachers. Accordingly, opportunities for professional development throughout teachers' careers offer a greater potential reach to improve teacher and school quality.

Australia, in general, has limited differences between graduate teachers' salaries and experienced teachers' salaries. The OECD (2016) reported that the top salary for lower secondary teachers was 44% percent higher than their starting salary, while across the OECD on average, the difference was 70%. Australian teachers also reach the top of their career progression more quickly relative to teachers in other countries – in just 8 years, compared to an OECD average of 25. This, arguably, discourages teachers from staying in the profession once they have reached their salary peak. As Ingvarson, Kleinhenz and Wilkinson (2007) put it: "...the implicit message in the salary scale is that teachers are not expected to improve their performance after nine years." This also gives high-quality teachers a financial incentive to pursue management or principal positions within schools, taking them out of classrooms.

In general, Australian teachers' salaries grow incrementally with each year of service. Each jurisdiction uses a different progression scale with its own salaries and steps. In addition, several jurisdictions offer an additional pay grade for teachers who have demonstrated higher levels of competency. This has historically been assessed against state based criteria.

More recently, the Highly Accomplished and Lead Teacher (HALT) levels of the Australian Professional Standards for Teachers (Teacher Standards) have aimed to provide a uniform standard for this assessment. HALT certification is a national initiative, implemented from 2013, which recognises skilled teachers and promotes the development of collaborative professionals who continually reflect upon and improve their practice in the classroom. There are currently 353 teachers (at December 2016) certified teachers across Australia in participating jurisdictions, from the government, independent and Catholic sectors in ACT, NT, SA, NSW and the independent sector in WA. Queensland is currently conducting a certification pilot project in 2017, with full implementation expected by 2019. Tasmania, Victoria and the government sector of WA do not currently participate in certification.

As noted above, Australia has a well-established workforce of teachers with an average of 16 years' experience. The impact of recent changes, like the HALT certification initiative, will take some time to appear; it may be many years before changes lead to changes at a classroom level.

Salaries and industrial arrangements more broadly for teachers have varied very little between jurisdictions historically. In any case, a top-down analysis of the entire school systems of states may not recognise the benefits of rewarding high-quality teachers. As Ingvarson, Kleinhenz and Wilkinson (2007) note:

While the impact that these teachers have on their students is likely to be significant, the impact that these schemes have in a wider sense across schools and school systems is probably small as the numbers of teachers in these positions is quite small.

There is insufficient evidence to draw a clear link from PISA and TIMSS performance of each jurisdiction's cohort to the incentives in industrial arrangements. As noted by Ingvarson, Kleinhenz and Wilkinson (2007), there has historically been relatively little variation between jurisdictions. Some recent changes to industrial agreements for teachers may change this: NSW is to have just 6 steps in its newest teacher agreement, with the earliest steps taking 2 years to complete (NSW Government, 2017); Victoria's pay scale has grown to 15 steps (Australian Education Union, 2017).

It is also noteworthy that there is little difference in outcomes between government and non-government schools once contextual factors have been controlled for (see section 4.1 above). Independent and Catholic schools have their own industrial arrangements with teachers outside of the enterprise bargaining agreements established by state and territory education departments.

Taken together, this policy analysis suggests that differences in industrial arrangements do not significantly drive school quality. Nonetheless, it is important to ensure that teachers are given incentives to continually develop throughout their careers. Effective incentives can drive improvements in school quality through levers that PISA and TIMSS indicate are more important but are harder to assess at a system level, such as teaching practice.

Teacher education course accreditation and teachers standards

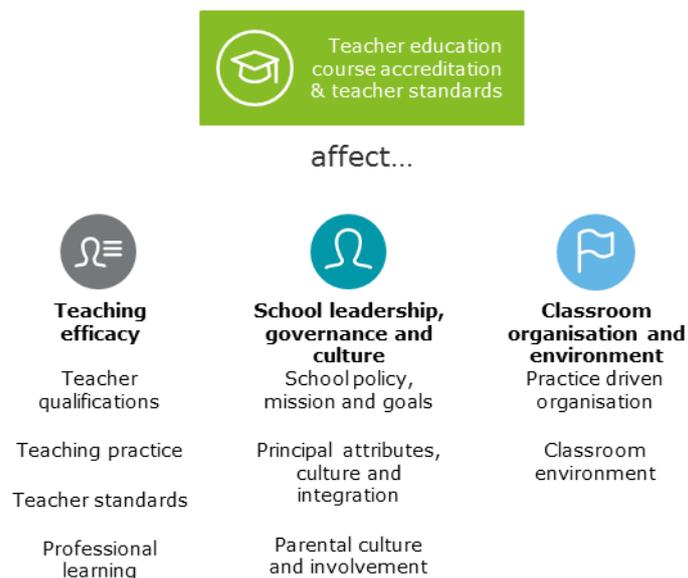
Another important role that government plays in determining the efficacy of the teaching workforce is in setting standards for teachers. There are two key standards of interest in the Australian context:

- **Pre-service teacher course accreditation** ensures that teachers enter the workforce with the necessary skills and attributes. The process aims to make universities accountable for the courses that they provide, recognising the important social role played by teachers.
- **Teacher professional standards** set a benchmark for the common capabilities that all teachers must possess across general teaching practice and articulate the range of skills all teachers should have to be effective in the classroom.

Beyond training teachers in evidence-based practices to improve student outcomes, there is increasing pressure on universities to focus on developing teachers' capacity to deal with behavioural issues, interacting with parents, and other skills necessary to render them "classroom ready" (Teacher Education Ministerial Advisory Group, 2014). The importance of this focus is emphasised by the relative importance of the classroom environment in influencing student outcomes, as outlined in section 4.2.1 above.

Much like industrial arrangements, these two forms of accreditation work across domains of school quality but ultimately provide an *indirect* lever to government to facilitate improving teacher attributes and practice (see Figure D.9).

Figure D.9: Impact of teacher course accreditation & teacher standards on school quality



Source: Deloitte Access Economics

Prior to 2011, each state and territory regulatory authority set its own requirements for the accreditation of initial teacher education courses. In 2011, all states and territories agreed to national accreditation standards. The agreed transition arrangements meant that courses were assessed for accreditation under the new national standards when their existing accreditation expired. By 2015, while many programs were accredited under the national accreditation standards, there were still a large number of programs accredited under the previous state and territory standards.

In December 2015, all states and territories endorsed the revised *Accreditation of initial teacher education programs in Australia: Standards and Procedures*. The development of the revised standards was led by the Australian Institute for Teaching and School Leadership (AITSL) as part of the Australian Government response to the Teacher Education Ministerial Advisory Group. State and territory regulatory authorities remain responsible for accrediting initial teacher education programs using these standards. It is expected that most initial teacher education programs will be accredited under the revised standards by the end of 2017.

In 2010 the Australian Government established the Australian Institute for Teaching and School Leadership (AITSL). In 2011 all Education Ministers endorsed the *Australian Professional Standards for Teachers* (Teacher Standards).

The Teacher Standards provide a nationally agreed quality assurance mechanism that ensures Australian teachers have the required competencies to be effective educators. They reflect the common capabilities that all teachers must possess and are designed to reflect quality across general teaching practice and to articulate the range of skills all teachers should have to be effective in the classroom. Although, ultimately, the responsibility of evaluating teachers falls to schools in line with the relevant industrial agreements, the Teacher Standards serve as a public statement of what constitutes classroom quality, and define what teachers should know and be able to do at different stages across their careers (Graduate, Proficient, Highly Accomplished and Lead). AITSL sets out three key stages of the ongoing cycle of assessing teacher performance (University of Melbourne Graduate School of Education, 2017):

- Reflection and goal setting;
- Professional practice and learning (information gathering); and
- Feedback and review.

In practice, each state and territory’s enterprise bargaining agreements vary in their approach to enforcing this evaluation framework. The University of Melbourne Graduate School of Education (2015) evaluated the Teacher Standards through a series of case studies. It noted that although the Standards provided important strategic and practical direction to systems, schools and teachers for ongoing development, they were used for different purposes at different schools. This would suggest that the change to a uniform standard of evaluation will require further time, and ongoing strategic direction, to result in improved outcomes at a student level.

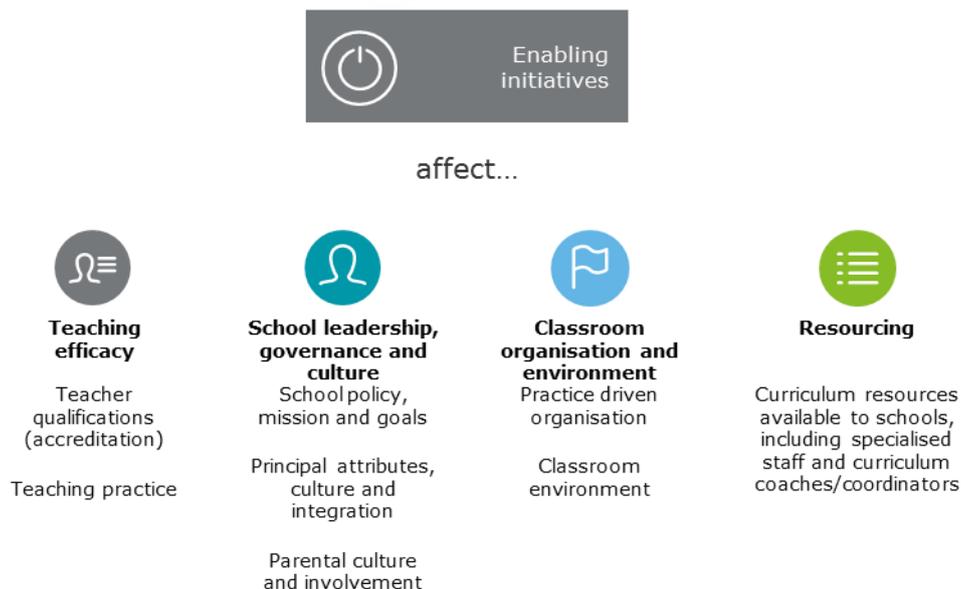
Enabling initiatives

Governments only directly determine the *framework* in which schools operate – the “system settings” described above. Ultimately, the practices of schools in classrooms are determined by teachers and leaders in those schools. Governments, however, have several tools which can be used to change that classroom practice:

- They can offer professional learning to schools;
- They can seek schools to participate in system-administered subject-specific initiatives;
- They can offer incentives for particular changes in practice at a school level.

These initiatives must necessarily be implemented by schools, rather than a system level—they rely on schools determining their relative areas of need and seeking interventions to implement.

Figure D.10: Overview of enabling initiatives



Source: Deloitte Access Economics

Teacher professional learning

Professional development—that is, “deliberate processes designed for the purposes of teacher post-initial professionally related education and training” (McRae et al., 2001)—provides a crucial link between the ever-expanding evidence base for effective interventions, and teacher implementation of those interventions. Recognising its overall importance, each state and territory government offers professional learning on an ongoing basis. Schools, in general, have capacity to select from a variety of professional learning opportunities from both public and private providers (McRae et al., 2001).

Teacher professional learning is deemed to be high quality when it includes opportunities for active learning and interaction with colleagues, is for an extended time period and comprises collective learning activities (e.g. communities of practice) or research with other teachers (OECD, 2017; B.

Jensen, 2014 and F. Barrera-Pedemonte, 2016). Professional learning is most effective when it is relevant, collaborative and future focused, and when it supports teachers to reflect on, question and consciously improve their practice (AITSL, 2012). Attendance at short-term conferences and seminars, workshops, discussions, lectures and field trips to other schools, are reported to have less impact on professional learning (AITSL, 2012).

Jurisdictions vary in their professional development programs in:

- The function, placement and size of the central organisation responsible for professional development;
- The nature of 'whole system' efforts as opposed to 'school initiated' efforts;
- The various requirement/entitlement structures in operation for both schools and teachers as individual employees; and
- The degree of autonomy in program selection at school level (McRae et al., 2001).

However, the literature notes that "evidence of the links between teacher professional development and student learning outcomes is hard to pin down" (Meiers and Ingvarson, 2005). The content of professional development programs significantly vary. Many may focus on areas outside of those measured by the outcomes recorded in PISA and TIMSS. Accordingly, its benefits may not be readily apparent in system-level data like that used in this analysis. Further detail on the role of government and schools in evaluating teacher professional learning is elaborated in Section 5.

Pedagogy initiatives

Each jurisdiction, at different times, has implemented initiatives aiming to develop teachers' approaches to particular areas of curriculum. In some cases, these initiatives are part of broader strategies which cut across several domains of school improvement: for instance, they may combine pedagogical changes with new accountability measures or increased resourcing. Background research conducted by the Department and Deloitte Access Economics has identified a number of such initiatives. Some case studies are included below. They are not intended to reflect current or best practice, but rather, to demonstrate the role governments and schools have historically played in such initiatives.

Western Australia – Getting it Right – Literacy and Numeracy Strategy

The Getting it Right – Literacy and Numeracy Strategy began in 2002 in Western Australia. It aimed to develop the expertise of teachers in order to improve outcomes and opportunities for children with literacy and numeracy difficulties. The Western Australian Department of Education placed Specialist Teachers in schools to focus on either literacy or numeracy. These teachers mentored and supported their colleagues and modelled effective teaching strategies. Principals of participating schools also set targets for measurable improvements in literacy and numeracy outcomes.

Meiers et al (2008) evaluated the impact of the program, using a combination of survey results and observations of classrooms and teachers. Teachers incorporated strategies from Specialist Teachers, and as a result, reported better belief in their self-efficacy and improved processes for identifying and teaching students at risk. Importantly, many teachers reported that they more explicitly identified goals and considered whether they had achieved them.

Tasmania – Raising the Bar

Raising the Bar, Closing the Gap, as it was originally known, began in 2008. Its aim is to increase the number of students completing primary school with functional literacy skills. It began in 36 regional primary schools, involving approximately 150 teachers and 2000 students from years 1 to 6 identified as being at or below the national minimum standard for literacy. The program provided participating schools with an additional assistant principal, releasing the principal to lead literacy and numeracy improvements within their schools. Participating schools also worked with a full-time literacy leader to facilitate discussions about improving literacy levels.

Hay et al (2011) evaluated the performance of the program using a combination of principal surveys and performance data from schools. They observed improvements in average scores across a number of standardised literacy tests. Principals also reported greater reliance on evidence for literacy planning.

Noting the positive results from evaluations of these programs, it is still difficult to identify their impacts at a system-wide level using PISA or TIMSS results:

- Participating schools run these programs in tandem with other State and Commonwealth programs.
- The programs are often run at select schools, particularly those below minimum literacy and numeracy standards. For instance, 365 schools, or around 50% of Western Australian government schools, participated in the Getting it Right program in the years studied by Meiers et al (2008).
- The programs often focus on a cohort of students within schools. For instance, although it aimed to improve literacy and numeracy achievement across all groups of students, Getting it Right focused on "Aboriginal students, students with a language background other than English (LBOTE), boys and students in rural and remote locations" – students who were more likely to face lower literacy and numeracy scores. Each participating school had one Specialist Teacher, who could be placed at the principal's discretion. Accordingly, a program may only have a real impact on student outcomes for a select group of students who were not participating in PISA or TIMSS.

Nonetheless, these types of targeted initiatives serve as a useful tool for improving student outcomes:

- They bring together a number of drivers of school quality – like improved accountability and teacher professional learning as a uniform intervention.
- They help to develop an improvement-focused culture at a school level. In particular, as part of a school-wide intervention, teachers are keen to demonstrate positive outcomes.

School leadership and management

School leaders can improve teaching and learning indirectly, through their influence on staff motivation, commitment and working conditions. Although many forms of school leadership are important in improving educational outcomes, those intended to influence pedagogy and which can be enacted by teacher leaders rather than principals are evidenced to be the most effective (Marks and Printy, 2003).

By cultivating a results-focused culture, and seeking evidence-based strategies to pursue system-wide goals, principals and senior teachers can provide school-level incentives to change classroom practice. Below, two case studies of leadership development initiatives in Western Australia and Victoria are outlined.

Western Australia – principal preparation programs

Anderson et al (2007) note that the career path of an Australian school principal follows an 'apprenticeship' model. All principals begin as teachers, gradually gain experience in leadership roles, and become principals over time. Although many aspiring principals pursue higher qualifications, they are not a requirement in Australian government school systems.

In Western Australia, a need was identified for further support for incoming principals. As a result, the Western Australian Department of Education's Leadership Centre established a number of training programs for aspiring principals.

For teachers and classroom leaders aspiring to principal roles, an online program *Explore: A Career in School Leadership* is available online. It is structured around the Australian Professional Standard for Principals, evaluating participants on their ability to incorporate those standards into their practice and develop a strong leadership vision and school culture.

In order to become formally eligible to apply for principal roles, candidates must complete a series of Principal Eligibility Modules. Taught by a school leader, the course helps participants to develop the necessary knowledge and skills in:

- Managing student services,
 - Managing workforce,
 - Managing financial resources, and
 - Managing risk.
-

Bastow Institute

The Bastow Institute of Educational Leadership, established in 2009, aims to “build the capacity of educational leaders and to identify and develop high quality leaders for the future.” Unlike professional learning institutes in other jurisdictions, it is specifically aimed at leadership training.

Its offerings include:

- A series of career stage programs for emerging leaders, middle leaders, aspiring and new principals, which provide structured professional learning that complement school practice;
- Courses on extending the capability of all education and early childhood education;
- Professional practice workshops of one to three days in length; and
- A number of thought leadership initiatives which focus in more depth on issues and trends in current learning theories.

These all provide school leaders with the opportunity to see exemplary practice modelled in action, and apply high-quality evidence-based practice in their own context. Thee professional learning programs at Bastow emphasise the role of school leaders in creating a school-wide focus on high performance, which acts at all levels to change classroom practice.

Limitation of our work

General use restriction

This report is prepared solely for the use of the Department of Education and Training. This report is not intended to and should not be used or relied upon by anyone else and we accept no duty of care to any other person or entity. The report has been prepared for the purpose of undertaking original empirical analysis and qualitative research into the drivers of school quality in Australia. You should not refer to or use our name or the advice for any other purpose.

Deloitte.

Access Economics

Contact us

Deloitte Access Economics

ACN: 149 633 116
8 Brindabella Circuit
Brindabella Business Park
Canberra Airport ACT 2609
Tel: +61 2 6263 7000
Fax: +61 2 6263 7004

Deloitte Access Economics is Australia's pre-eminent economics advisory practice and a member of Deloitte's global economics group. For more information, please visit our website

www.deloitte.com/au/deloitte-access-economics

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/au/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms.

The entity named herein is a legally separate and independent entity. In providing this document, the author only acts in the named capacity and does not act in any other capacity. Nothing in this document, nor any related attachments or communications or services, have any capacity to bind any other entity under the 'Deloitte' network of member firms (including those operating in Australia).

About Deloitte

Deloitte provides audit, tax, consulting, and financial advisory services to public and private clients spanning multiple industries. With a globally connected network of member firms in more than 150 countries, Deloitte brings world-class capabilities and high-quality service to clients, delivering the insights they need to address their most complex business challenges. Deloitte's approximately 200,000 professionals are committed to becoming the standard of excellence.

About Deloitte Australia

In Australia, the member firm is the Australian partnership of Deloitte Touche Tohmatsu. As one of Australia's leading professional services firms, Deloitte Touche Tohmatsu and its affiliates provide audit, tax, consulting, and financial advisory services through approximately 6000 people across the country. Focused on the creation of value and growth, and known as an employer of choice for innovative human resources programs, we are dedicated to helping our clients and our people excel. For more information, please visit our web site at www.deloitte.com.au.

Liability limited by a scheme approved under Professional Standards Legislation.

Member of Deloitte Touche Tohmatsu Limited